The Identification Power of Equilibrium in Simple Games*

Andres Aradillas-Lopez†
Princeton University

Elie Tamer‡
Northwestern University

February 2008

Abstract

We examine the identification power that (Nash) equilibrium assumptions play in conducting inference about parameters in some simple games. We focus on three static games where we drop the Nash equilibrium assumption and instead use *rationalizability* (Bernheim (1984) and Pearce (1984)) as the basis for strategic play. The first example examines a bivariate discrete game with complete information of the kind studied in entry models. The second example considers the incomplete information version of the discrete bivariate game. Finally, the third example considers a first price auction with independent private values. In each example, we study the inferential question of what can be learned about the parameter of interest using a random sample of observations, under *level-k rationality* where k is an integer ≥ 1. As k increases, our identified set shrinks, limiting to the identified set under full rationality or rationalizability (as $k \to \infty$). This is related to the concept of *higher order beliefs*, which are incorporated into the econometric analysis in our framework. We are then able to categorize what can be learned about the parameters in a model under various maintained levels of rationality, highlighting the role different assumptions play. We provide constructive identification results that lead naturally to consistent estimators.

*We thank seminar participants many institutions for comments. We also thank M. Engers, D. Fudenberg, F. Molinari, M. Sinischalchi, and A. Pakes for useful comments. We especially thank Serena Ng for her diligent editorial help and the discussants for their valuable comments.

†Department of Economics, Princeton University, Princeton NJ 08544. Support from the National Science Foundation and from the Gregory C. Chow Econometric Research Program at Princeton University are gratefully acknowledged.

‡Department of Economics, Northwestern University, Evanston IL 60208. Support from the National Science Foundation is gratefully acknowledged.

1 Introduction

This paper examines the identification power of equilibrium in some simple games. In particular, we relax the assumption of Nash Equilibrium (NE) behavior and assume that players are rational. Rationality posits that agents play strategies that are consistent with a set of proper beliefs. The object of interest in these games is a parameter vector that parametrizes payoff functions. We compare what can be learned about this vector from a random sample of data, under a set of rationality assumptions, culminating with rationalizability, a concept introduced jointly in the literature by Bernheim (1984) and Pearce (1984) . We find that in static discrete games with *complete information*, the identified features of the games with more than one level of rationality is similar to what one obtains with Nash behavior assumption but allowing for multiple equilibria (including equilibria in mixed strategies). In a bivariate game with *incomplete information,* if the game has a unique (Bayesian) Nash Equilibrium, then there is convergence between the identified features with and without equilibrium only when the level of rationality tends to infinity. When there is multiple equilibria, the identified features of the game under rationalizability and equilibrium are different: smaller identified sets (hence more information about the parameter of interest) when equilibrium is imposed, but computationally easier to construct identification regions when imposing rationalizability (no need to solve for fixed points). In the auction game we study, the situation is different. We follow the work of Battigalli and Sinischalchi (2003) where under some assumptions, given the valuations, rationalizability predicts only upper bounds on the bids. We show how these bounds can be used to learn about learn about the latent distribution of valuation. Another strategic assumptions in auctions resulting in tighter bounds is the concept of $\mathcal{P}$-dominance studied in Dekel and Wolinsky (2003).

Economists have observed that equilibrium play in noncooperative strategic environment is not necessary for rational behavior. Some can easily construct games where NE strategy profiles are unreasonable while on the other hand, one can also find reasonable strategy profiles that are not Nash. Restrictions once Nash behavior is dropped are based typically on a set of "rationality" criteria. These criteria are enumerated in different papers and under different strategic scenarios. This paper studies the effect of adopting a particular rationality criterion on learning about parameters of interests. We do not advocate one type of strategic assumption over another, but simply explore one alternative to Nash and see its effect on parameter inference. Thus, depending on the application, identification of parameters of interest can certainly be studied under strategic assumptions other than rationalizability. We provide such an example in this paper.

Since every Nash profile is rational under our definition, dropping equilibrium play complicates the identification problem because under rationality only, the set of predictions is enlarged. As Pearce notes, "this indeterminacy is an accurate reflection of the difficult situation faced by players in a game" since logical guidelines and the rules of the game are not sufficient for uniqueness of predicted behavior. Hence, it is interesting from the econometric perspective to examine how the identified features of a particular game changes as weaker assumptions on behavior are made.

We maintain that players in the game are rational where heuristically, we define rationality as behavior that is consistent with an optimizing agent equipped with a proper set of beliefs (or probability distributions) about the unknown actions of others. Rationality comes in different levels or orders where a profile is first order rational (or rationalizable) if it is a best response to some profile for the other players. This intersection of layers of rationality constitutes rationalizable strategies. We study the identification question for level-k rationality for $k \geq 1$. When we study the identifying power of a game under a certain set of assumptions on the strategic environment, we implicitly assume that all players in that game are abiding exactly by these assumptions and playing exactly that game[1].

Using equilibrium as a restriction to gain identifying power is well known in economics[2]. The objective of the paper is to study the identification question in (simple) game-theoretic models without the assumption of equilibrium– by focusing on the weaker concept of rationality –k-level rationality and its limit rationalizability– of strategies and beliefs. This approach has two important advantages. First, it leads naturally to a well-defined concept of *levels of rationality* which is attractive practically. Second, it can be adapted to a very wide class of models without the need to introduce ad-hoc assumptions. Ultimately, interim rationalizability allows us to do inference (to varying degrees) both on the structural parameters of a model (for example, the payoff parameters in a reduced-form game, or the distribution of valuations in an auction), as well as on the properties of *higher order beliefs* by the agents, which are incorporated into the econometric analysis. The features of this hierarchy of beliefs will characterize what we refer to as the *rationality-level* of agents. In addition, it is possible to also provide testable restrictions that can be used to find an upper bound on the rationality level in a give data set.

Level-k thinking as an alternative to Nash equilibrium behavior has also been studied in Stahl and Wilson (1995), Nagel (1995), Ho, Camerer, and Weigelt (1998), Costa-Gomes, Crawford, and Broseta (2001), Costa-Gomes and Crawford (2006) and Crawford and Iriberri (2007a). These models depart from equilibrium behavior by dropping the assumption that each player has a perfect model of others' decisions and replacing it with the assumption that such subjective models survive k rounds of iterated elimination of dominated decisions. Thus, each player's subjective model about others' behavior is consistent with level-k interim rationalizability in the sense of Bernheim (1984). Their identification strategy is to assume the existence of a small number of pre-specified types, each of which is associated with a very specific behavior (for example, a particular type of player could perform two mental rounds of deletion of dominated strategies and best-respond to a uniform distribution over the surviving actions). Using carefully designed experiments, these researchers have sought to explain which type fits the observed choices the best. Our work differs fundamentally from the aforementioned papers because we focus on identification based exclusively on the bounds for conditional choice probability that result from k-steps of deletion of dominated strategies in a semiparametric model. Even though such bounds are valid for a wide number of "behavioral types", we do not focus on any single one of them.ANDRES: can you clarify this? In addition, we focus on situations where the researcher ignores how "rational" players are and where other primitives of the game are also the object of interest: Payoff parameters in discrete games, or the distribution of valuations in an auction. In an experimental data set, the last set of objects are entirely under the control of the researcher, and strong parametric assumptions are typically made about behavioral types.

In the first section, we review and define rational play in a noncooperative strategic game. Here we mainly adapt the definition provided in Pearce. We then examine the identification power of dropping Nash behavior in some commonly studied games in empirical economics. In section 3, we consider discrete static games of complete information. This type of game is widely used in the empirical literature on (static) entry games with complete information and under Nash equilibrium (See Bjorn and Vuong (1985), Bresnahan and Reiss (1991), Berry (1994), Tamer (2003), Andrews, Berry, and Jia (2003), Ciliberto and Tamer (2003), and Bajari, Hong, and Ryan (2005) among others). Here, we find that in the 2×2 game with level-2 rationality, the outcomes of the game coincide with Nash, and hence econometric restrictions are the same. Section 4 considers static games with incomplete information. Empirical frameworks for these games are studied in Aradillas-Lopez (2005), Aguiregabiria and Mira (2004), Seim (2002), Pakes, Porter, Ho, and Ishii (2005), Berry and Tamer (2006) among others. Characterization of rationalizability in the incomplete information game is closely related to the higher-order belief analysis in the *global games* literature (see Morris and Shin (2003)) and to other recently developed concepts such as those in Dekel, Fudenberg, and Morris (2007) and Dekel, Fudenberg, and Levine (2004). Here, we show that level-k rationality implies restrictions on player beliefs in the 2×2 game that lead to simple restrictions that can be exploited in identification. As k increases, an iterative elimination procedure restricts the size of the allowable beliefs which map into stronger restrictions that can be used for identification. If the game admits a unique equilibrium, the restrictions of the model converge towards Nash restrictions as the level of rationality k increases. With multiple equilibria, the iterative procedure converges to sets of beliefs that contain both the "large" and "small" equilibria. In particular, studying identification in these settings is simple since one does not need to solve for fixed points, but to simply iterate the beliefs towards the predetermined level of rationality k. In section 5 we examine a first price independent auction game where we follow the work of Battigalli and Sinischalchi (2003). Here, for any order k, we are only able to bound the valuation from above. Finally, Section 6 concludes.

2 Nash Equilibrium and Rationality

In noncooperative strategic environment, optimizing agents maximize a utility function that depends on what their opponents do. In simultaneous games, agents attempt to predict what their opponents will play, and then play accordingly. Nash behavior posits that players' expectations of what others are doing are mutually consistent, and so a strategy profile is Nash if no player has an incentive to change their strategy given what the other agents are playing. This Nash behavior makes an implicit assumption on players' expectations. But, players "are not compelled by deductive logic" (Bernheim) to play Nash. In this paper, we examine the effect of assuming Nash behavior on identification by comparing restrictions under Nash with ones obtained under rationality in the sense of Bernheim and Pearce. Below, we follow Pearce's framework and first maintain the following assumptions on behavior:

- Players use proper subjective probability distribution, or use the axioms of Savage, when analyzing uncertain events.
- Players are expected utility maximizers.
- Rules and structure of the game are common knowledge.

We next describe heuristically what is meant by rationalizable strategies. Precise definitions are given in Pearce (1984) for example.

- We say that a strategy profile for player i (which can be a mixed strategy) is **dominated** if there exists another strategy for that player that does better no matter what other agents are playing.

- Given a profile of strategies for all players, a strategy for player i is a **best response** if that strategy does better for that player than any other strategy given that profile.

To define rationality, we make use of the following notations. Let $\mathcal{R}^i(0)$ be the set of all (possibly mixed) strategies that player i can play and $\mathcal{R}^{-i}(0)$ is the set of all strategies for players other than i. Then, heuristically:

- **Level-1 rational strategies** for player i are strategy profiles $s^i \in \mathcal{R}^i(0)$ such that there exists a strategy profile for other players in $\mathcal{R}^{-i}(0)$ for which s^i is a best response. The set of level-1 strategies for player i is $\mathcal{R}^i(1)$.

- **Level-2 rational strategies** for player i are strategy profiles $s^i \in \mathcal{R}^i(0)$ such that there exists a strategy profile for other players in $\mathcal{R}^{-i}(1)$ for which s^i is a best response.

- **Level-t rational strategies:** Defined recursively from level 1.

Notice that by construction, $\mathcal{R}^i(1) \subseteq \mathcal{R}^i(0)$ and $\mathcal{R}^i(t) \subseteq \mathcal{R}^i(t-1)$. Finally, **rationalizable strategies** are ones that lie in the intersection of the $\mathcal{R}$'s as t increases to infinity. Moreover, one can show (See Pearce) that there exists a finite k such that for $\mathcal{R}^i(t) = \mathcal{R}^i(k)$ for all $t \geq k$.

Rationality in these settings is equivalent to best response, in that, a strategy is rational for a player if it is a best response to some strategy profile by other players. If we iterate this further, we arrive at the set of rationalizable strategies. Pearce provided properties of the rationalizable set. For example, NE profiles are always included in this set and this set contains at least one profile in pure strategies.

3 Bivariate Discrete Game with Complete Information

Consider the following bivariate discrete 0/1 game where t_p is the payoff that player p obtains by playing 1 when player $-p$ is playing 0. We have parameters α_1 and α_2 that are of interest. The econometrician does not observe t_1 or t_2 and is interested in learning about the α's and the joint distribution of (t_1, t_2). Assume also, as in entry games, that the α's are negative. In this example and the next, we assume that one has access to a random

Table 1: Bivariate Discrete Game

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	0,0	0,t_2
$a_1 = 1$	t_1,0	$t_1 + \alpha_1,\ t_2 + \alpha_2$

sample of observations $(y_{1i}, y_{2i})_{i=1}^N$ which represent for example market structures in a set of N independent markets. To learn about the parameters, we map the observed distribution of the data (the choice probabilities) to the distribution (or set of distributions) predicted by the model. Since this is a game of complete information, players observe all the payoff relevant information. In particular, in the first round of rationality, player 1 will play 1 if $t_1 + \alpha_1 \geq 0$ since this will be a dominant strategy. In addition, if t_1 is negative, player 1 will play 0. However, when $t_1 + \alpha_1 \leq 0 \leq t_1$, both actions 1 and 0 are level-1 rational: action 1 is rational since it can be a best response to player 2 playing 0, while action 0 is a best response to player 2 playing 1. The set $\mathcal{R}(1)$ is summarized in Figure 1 below. For example, consider the upper right hand corner. For values of t_1 and t_2 lying there, playing 0 is not a best response for either player. Hence, $(1, 1)$ is the unique level-1 rationalizable strategy (which is also the unique NE). Consider now the middle region on the right hand side, i.e., $(t_1, t_2) \in [-\alpha_1, \infty) \times [0, -\alpha_2]$. In level-1 rationality, 0 is not a best reply for player 1, but 2 can play either 1 or 0 : 1 is a best reply when 1 plays 0, and 0 is a best reply for player 2 when 1 plays 1. However, in the next round of rational play, given that player 2 now believes that player 1 will play 1 with probability 1, then player 2's response is to play 0. Hence $\mathcal{R}(1) = \{\{1\}, \{0, 1\}\}$ while the rationalizable set reduces to the outcome $(1, 0)$. Here, $\mathcal{R}(k) = \mathcal{R}(2) = \{\{1\}, \{0\}\}$ for all $k \geq 2$. In the middle square, we see that the game provides no observable restrictions: any outcome can be *potentially observable* since both strategies are rational at any level of rationality. Notice also that in this game, the set of rationalizable strategies is the set of profiles that are undominated. This is a property of

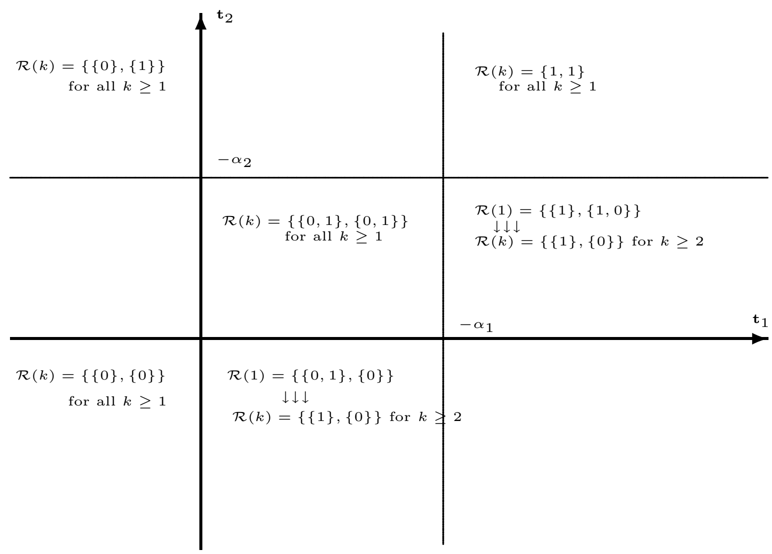


Figure 1: Rationalizable Profiles in a Bivariate Game with Complete Information

bivariate binary games.

3.1 Infence with level-k rationality:

A random sample of observations allows us to obtain a consistent estimator of the choice probabilities (or the data). The object of interest here is $\theta = (\alpha_1, \alpha_2, F(.,.))$ where $F(.,.)$ is the joint distribution of (t_1, t_2). One interesting approach to conduct inference on the sharp set is to assume that both t_1 and t_2 are discrete random variables with identical support on $s_1, \ldots, s_K$ such that $P(t_1 = s_i; t_2 = s_j) = p_{ij} \geq 0$ for $i, j \in \{1, \ldots, k\}$ with $\sum_{i,j} p_{ij} = 1$. Hence, we make inference on on the set of probabilities $(p_{ij},\, i, j \leq k)$ and (α_1, α_2). We highlight this below for level 2 rationality. In particular, we say that

$$\theta = ((p_{ij}), \alpha_1, \alpha_2) \in \Theta_I$$

if and only if:

$$P_{11} = \sum_{i,j} p_{ij} \left(1[s_i \geq -\alpha_1;\; s_j \geq -\alpha_2] + l_{ij}^{(1,1)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right)$$
$$P_{00} = \sum_{i,j} p_{ij} \left(1[s_i \leq 0;\; s_j \leq 0] + l_{ij}^{(0,0)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right)$$
$$P_{10} = \sum_{i,j} p_{ij} \left(1[s_i \geq 0;\; s_j \leq 0] + 1[s_i \geq -\alpha_1; 0 \leq s_j \leq -\alpha_2] + l_{ij}^{(1,0)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right)$$
$$P_{01} = \sum_{i,j} p_{ij} \left(1[s_i \leq 0;\; s_j \geq 0] + 1[0 \leq s_i \leq -\alpha_1; s_j \geq -\alpha_2] + l_{ij}^{(0,1)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right)$$

for some $(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) \geq 0$ and $l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1$ for all $i, j \leq k$. One can think of the l's as the "selection mechanisms" that pick an outcome in the region where the model predicts multiple outcomes. We are treating the support points as known, but this is without loss of generality since those too can be made part of θ. The above equalities (and inequalities), for a given θ, are similar to first order conditions from a linear programming problem and hence can be solved fast using linear programming algorithms. In particular, Consider the objective function in (3.1) below. Note first that $Q(\theta) \leq 0$ for all θ in the parameter space. And,

$$\theta \in \Theta_I$$

if and only if $Q(\theta) = 0$.

$$\begin{aligned}
Q(\theta) = & \max_{v_i,\ldots,v_8,(l_{ij}^{(1,1)},l_{ij}^{(0,0)},l_{ij}^{(0,1)},l_{ij}^{(1,0)})} -(v_1 + \ldots, +v_8) \qquad s.t. \\
& P_{11} - \sum_{i,j} p_{ij} \left(1[s_i \geq -\alpha_1;\; s_j \geq -\alpha_2] + l_{ij}^{(1,1)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right) = v_1 - v_2 \\
& P_{00} - \sum_{i,j} p_{ij} \left(1[s_i \leq 0;\; s_j \leq 0] + l_{ij}^{(0,0)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right) = v_3 - v_4 \\
& P_{10} - \sum_{i,j} p_{ij} \left(1[s_i \geq 0;\; s_j \leq 0] + 1[s_i \geq -\alpha_1; 0 \leq s_j \leq -\alpha_2] + l_{ij}^{(1,0)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right) = v_5 - v_6 \\
& P_{01} - \sum_{i,j} p_{ij} \left(1[s_i \leq 0;\; s_j \geq 0] + 1[0 \leq s_i \leq -\alpha_1; s_j \geq -\alpha_2] + l_{ij}^{(0,1)} 1[0 \leq s_i \leq -\alpha_1;\; 0 \leq s_j \leq -\alpha_2]\right) = v_7 - v_8 \\
& v_i \geq 0; \quad (l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) \geq 0; \quad l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1 \quad \text{for all}\, 1 \leq i, j \leq k
\end{aligned} \tag{3.1}$$

First, note that for any θ, the program is feasible: for example, set $(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) = 0$ and then set $v_1 = P_{11} - \sum_{i,j} p_{ij} 1[s_i \geq -\alpha_1;\; s_j \geq -\alpha_2]$ and $v_2 = 0$ if $P_{11} - \sum_{i,j} p_{ij} 1[s_i \geq -\alpha_1;\; s_j \geq -\alpha_2] \geq 0$, otherwise set $v_2 = -(P_{11} - \sum_{i,j} p_{ij} 1[s_i \geq -\alpha_1;\; s_j \geq -\alpha_2])$ and $v_1 = 0$ and similarly for the rest. Moreover, $\theta \in \Theta_I$ if and only if $Q(\theta) = 0$. One can collect all the parameter values for which the above objective function is equal to zero (or approximately equal to zero) A similar linear programming procedure was used in Honoré and

Tamer (2005). The sampling variation comes from having to replace the choice probabilities $(P_{11}, P_{12}, P_{21}, P_{22})$ with their sample analogs which would result in a sample objective function $Q_n(.)$ that can be used to conduct inference.

More generally, and without making support assumptions, a practical way to conduct inference if assumes one level of rationality say, is to use an implication of the model. In particular, under $k = 1$ rationality, the statistical structure of the model is one of moment inequalities:

$$\begin{aligned}
\Pr(t_1 \geq -\alpha_1; t_2 \geq -\alpha_2) &\leq P(1,1) \leq \Pr(t_1 \geq 0; t_2 \geq 0) \\
\Pr(t_1 \leq 0; t_2 \leq 0) &\leq P(0,0) \leq \Pr(t_1 \leq -\alpha_1; t_2 \leq \alpha_2) \\
\Pr(t_1 \geq -\alpha_1; t_2 \leq 0) &\leq P(1,0) \leq \Pr(t_1 \geq 0; t_2 \leq -\alpha_2) \\
\Pr(t_1 \leq 0; t_2 \geq -\alpha_2) &\leq P(0,1) \leq \Pr(t_1 \leq -\alpha_1; t_2 \geq 0)
\end{aligned}$$

The above inequalities do not exploit all the information and hence the identified set based on these inequalities is not sharp[3]. However, these inequalities based moment conditions are simple to use and can be generalized to large games. Heuristically then, the model identifies, by definition, the set of parameters Θ_I such that the above inequalities are satisfied. Moreover, we say that the model *point identifies* a unique θ if the set Θ_I is a singleton.
In the next figure, we provide the mapping between the predictions of the game and the observed data under Nash and level-k rationality. The observable implication of Nash is

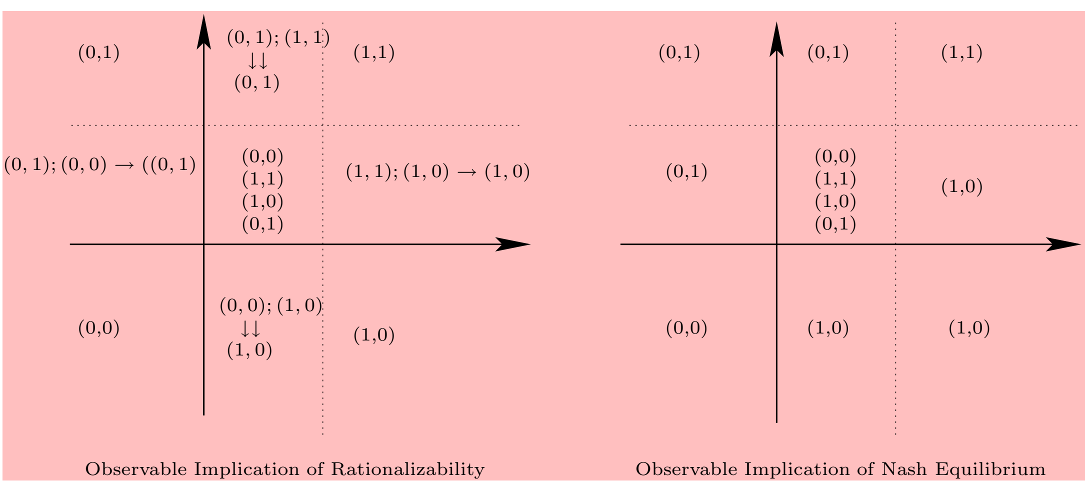
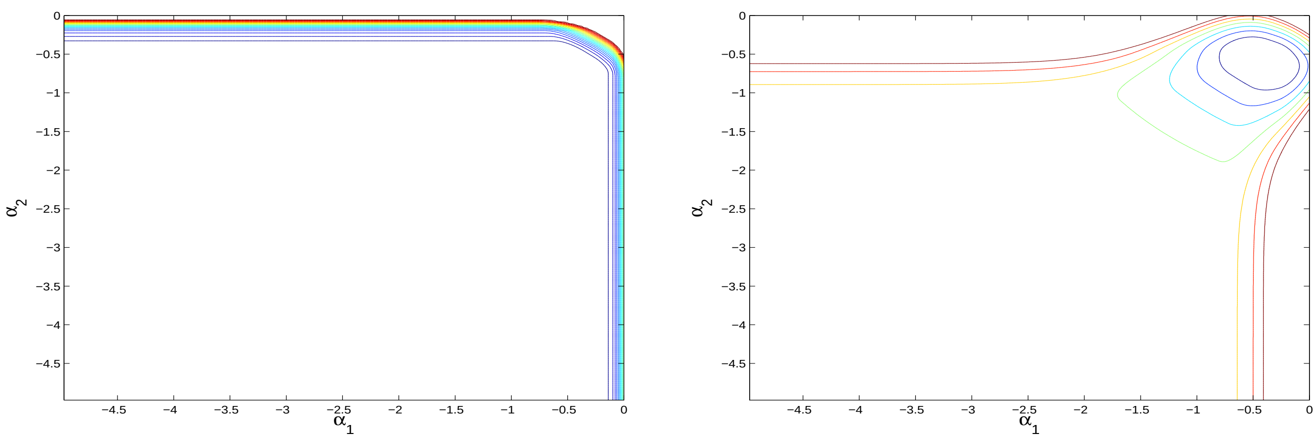


Figure 2: Observable Implications of Equilibrium vs Rationality

different depending on whether we allow for mixed strategies. In particular, without allowing for mixed strategies, in the middle square of the Nash figure, the only observable implication is $(1,0)$ and $(0,1)$. However, it reverts to all outcomes once one consider the mixed strategy equilibrium.

To a get an idea of the identification gains when we assume rationality vs equilibrium, we simulated a stylized version of the above game in the case where t_i is standard normal for $i = 1, 2$ and the only object of interest is the vector (α_1, α_2). We compare the identified set of the above game under $k = 1$ rationality and NE when we only consider pure strategies. We see from Figure 3 that there is identifying power in assuming Nash equilibrium. In

Figure 3: Identification Set under Nash and 1-level Rationality



Above we see the identified regions for (α_1, α_2) under $k = 1$-rationality (left display) and Nash (right display). We set in the underlying model $(\alpha_1, \alpha_2) = (-.5, -.5)$ (The model was simulated assuming Nash with $(0, 1)$ selected with probability one in regions of multiplicity.) Notice that on the left, the model only places upper bounds on the alphas. Under Nash on the other hand, (α_1, α_2) are constrained to lie a much smaller set (the inner "circle").

particular, under Nash, the identified set is a somehow tight "circle" around the simulated truth while under rationality, the model only provides upper bounds on the alpha's. But, if we add exogenous variations in the profits (X's), the identified region under rationality will shrink. The next section examines the identifying power of the same game under incomplete information.

4 Discrete Game with Incomplete Information

Consider now the discrete game presented in Table 1 above but under the assumptions that player 1 (2) does not observe t_2 (t_1) or that the signals are private information. We will denote player $p \in \{1,2\}$'s opponent to be $-p$. Both players hold beliefs' about anothers' type and these beliefs can be summarized by a subjective distribution function. For player 1 of type t_1, let the space of beliefs about 2's probability of entry be $\mathbb{P}_{t_1} \equiv \Pr(a_2 = 1|t_1)$ which can depend on t_1. Given belief $\mathbb{P}_{t_1}$, the expected utility function of player 1 is

$$U(a_1, P_{t_1}) = \begin{cases} t_1 + \alpha_1 \mathbb{P}_{t_1} & \text{if } a_1 = 1 \\ 0 & \text{otherwise} \end{cases}$$

Similarly for player 2 we have

$$U(a_2, P_{t_2}) = \begin{cases} t_2 + \alpha_1 \mathbb{P}_{t_2}(a_1 = 1) & \text{if } a_2 = 1 \\ 0 & \text{otherwise} \end{cases}$$

Again, we assume here both α_1 and α_2 are negative.

4.1 Identification without assumptions on beliefs: *level*-1 rationality:

In the first round of rationality, we know that *for any belief function*, or without making any common prior assumptions, the following holds:

$$\begin{aligned} t_1 + \alpha_1 \geq 0 \quad &\Longrightarrow U(1, \mathbb{P}_{t_1}) = t_1 + \alpha_1 \mathbb{P}_{t_1} \geq 0 \quad \forall \mathbb{P}_{t_1} \in [0,1] \\ t_1 \leq 0 \quad &\Longrightarrow U(1, \mathbb{P}_{t_1}) = t_1 + \alpha_1 \mathbb{P}_{t_1} \leq 0 \quad \forall \mathbb{P}_{t_1} \in [0,1] \end{aligned} \tag{4.2}$$

which implies that

$$\begin{aligned} t_1 + \alpha_1 \geq 0 \quad &\Longrightarrow a_1 = 1 \\ t_1 \leq 0 \quad &\Longrightarrow a_1 = 0 \end{aligned}$$

This is common knowledge among the players. Now, let $0 \leq t_1 \leq -\alpha_1$. For a player that is rational of order one, there exists well defined beliefs that rationalizes either 1 or 0. Hence, when $0 \leq t_1 \leq -\alpha_1$, both $a_1 = 1$ and $a_1 = 0$ are rationalizable. So, the implication of the game are summarized in Figure 4 below. Notice here that the (t_1, t_2) space is divided into 9 regions: 4 regions where the outcome is unique, 4 regions with 2 potentially observable outcomes, and the middle square where any outcome is potentially observed. To make inference based on this model, one needs to map these regions into predicted choice probabilities. To obtain the sharp set of parameters that is identified by the model, one can supplement this model with consistent "selection rules" that specifies in regions of multiplicity, a well defined probability of observing the various outcomes (this would be a function of of both t_1 and t_2). Another practical approach to inference in these

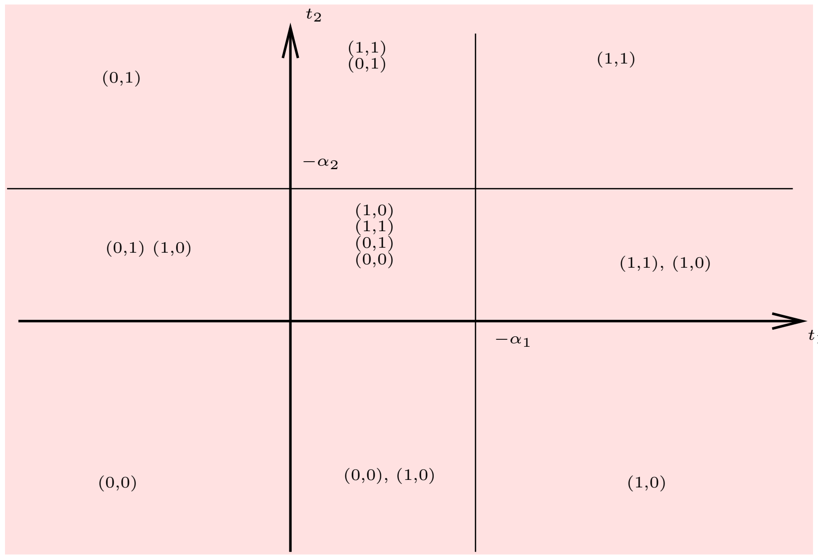


Figure 4: Observable Implications of Level-1 Rationality

settings is to exploit implication of the model, mainly that these selection function are probabilities and hence bounded between zero and one. These implications will imply the following restrictions in terms of moment inequalities[1]:

Result 1 *For the game with incomplete information, let the players be rational with order 1. Then the model implies the following predictions on choice probabilities:*

$$\begin{aligned} P(t_1+\alpha_1 \geq 0; t_2+\alpha_2 \geq 0) \leq P(1,1) \leq P(t_1 \geq 0; t_2 \geq 0) \\ P(t_1 \leq 0; t_2 \leq 0) \leq P(0,0) \leq P(t_1 \leq -\alpha_1; t_2 \leq -\alpha_2) \\ P(t_1 \leq 0; t_2 \geq -\alpha_2) \leq P(0,1) \leq P(t_1 \leq -\alpha_1; t_2 \geq 0) \\ P(t_1 \geq -\alpha_1; t_2 \leq 0) \leq P(1,0) \leq P(t_1 \geq 0; t_2 \leq 0) \end{aligned} \tag{4.3}$$

The above inequalities can be exploited in a similar way as we did in the previous section to construct the identified set. The latter, Θ_I, is the set of parameters for which the above inequalities in (4.3) are satisfied.

Now, we add another round of rationality. To do that, the beliefs of the players, though not required to be "correct", play a role. It matters what each player believes the other's likelihood of playing 1 or 0 is. To complete the description of this strategic situation, we make a common prior assumption. We expand on this below.

[1]Again, here, there are cross equation restrictions that can be used. So, ignoring these will result in larger identified sets.

4.2 Identification with *"level-k"* ($k > 1$) Rationalizable Beliefs

Here, we study the identification question with higher order rationality. As a reminder, player p with type t_p, though does not observe t_{-p}, but treats the latter as a random variable from a known (conditional) distribution $\mathbb{P}_{t_p}$. The form of this distribution is common knowledge to both players (i.e., we maintain a *joint* distribution on the types and this joint distribution is common knowledge). This is the common prior assumption. We also allow that this joint distribution be conditional on some information set for player p, call it $\mathcal{I}_p$ (which includes t_p). We will consider strategies for player p that are threshold functions of t_p:

$$Y_p = \mathbb{1}\{t_p \geq \mu_p\} \quad \text{for } p = 1, 2, \tag{4.4}$$

Therefore, player p forms *subjective* beliefs about μ_{-p} that can be summarized by a probability distribution for μ_{-p} given $\mathcal{I}_p$. These beliefs are derived as part of a solution concept. For example they include BNE beliefs as a special case (in which case all players know those equilibrium beliefs to be correct). Here, let $\widehat{G}_1(\mu_2|\mathcal{I}_1)$ denote Player 1's subjective distribution function for μ_2 given $\mathcal{I}_1$, and define $\widehat{G}_2(\mu_1|\mathcal{I}_2)$ analogously for player 2. A strategy by player p is *rationalizable* if it is the best response (in the expected-utility sense) given some beliefs $\widehat{G}_p(\mu_{-p}|\mathcal{I}_p)$ that assign zero probability mass to strictly dominated strategies by player $-p$. **A rationalizable strategy by player** p is described by

$$Y_p = \mathbb{1}\Big\{t_p + \alpha_p \int_{\mathbb{S}(\widehat{G}_p)} E\big[\mathbb{1}\{t_{-p} \geq \mu\}\big|\mathcal{I}_p, \mu\big] d\widehat{G}_p(\mu|I_p) \geq 0\Big\}, \tag{4.5}$$

where the support $\mathbb{S}(\widehat{G}_p)$ excludes values of μ that result in strictly dominated strategies within the class (4.4). Note that the subset of rationalizable strategies within the class (4.4) are of the form $\mu_p = -\alpha_p \int_{\mathbb{S}(\widehat{G}_p)} E\big[\mathbb{1}\{t_{-p} \geq \mu\}\big|\mathcal{I}_p, \mu\big] d\widehat{G}_p(\mu|I_p)$. In this setting, rationalizability requires expected utility maximization for a given set of beliefs, but it does not require those beliefs to be correct. It only imposes the condition that $\mathbb{S}(\widehat{G}_p)$ exclude values of μ_{-p} that "do not make sense" or are dominated. We eliminate such values by iterated deletion of dominated strategies.

Again, we maintain that the signs of the strategic-interaction parameters (α_1, α_2) are known. Specifically, suppose $\alpha_p \leq 0$. Then, repeating arguments from the previous section on k = 1-rationalizable outcomes, we see looking at (4.5) that we must have (event-wise comparisons):

$$\mathbb{1}\{t_p + \alpha_p \geq 0\} \leq \mathbb{1}\{Y_p = 1\} \quad \text{and} \quad \mathbb{1}\{t_p < 0\} \leq \mathbb{1}\{Y_p = 0\}$$

Decision rules that do not satisfy these conditions are strictly dominated for all possible beliefs. Therefore, the subset of strategies within the class (4.4) that are not strictly dominated must satisfy $\Pr(t_p + \alpha_p \geq 0) \leq \Pr(t_p \geq \mu_p) \leq \Pr(t_p \geq 0)$, or equivalently, $\mu_p \in [0, -\alpha_p]$. All other values of μ_p correspond to dominated strategies. We refer, in this set-up, to the subset of strategies that satisfy $\mu_p \in [0, -\alpha_p]$ as *level-1 rationalizable strategies.* Note, as before, that these μ's do NOT involve the common prior distributions.

Players are level-2 rational if it is common knowledge that they are level-1 rational and they use this information rationally. Level-2 rationalizable beliefs by player p must assign zero probability mass to values $\mu_{-p} \notin [0, -\alpha_p]$. A strategy is level-2 rationalizable if it can be justified by level-2 rationalizable beliefs. That is, a strategy $Y_p = \mathbb{1}\{t_p \geq \mu_p\}$ is *level-2 rationalizable* if

$$\mu_p = -\alpha_p \int_0^{-\alpha_{-p}} E\big[\mathbb{1}\{t_{-p} \geq \mu\}\big|\mathcal{I}_p, \mu\big] d\widehat{G}_p(\mu|\mathcal{I}_p)$$

where player p's beliefs $\widehat{G}_p(\cdot|\mathcal{I}_p)$ satisfy $\widehat{G}_p(0|\mathcal{I}_p) = 0$ and $\widehat{G}_p(-\alpha_{-p}|\mathcal{I}_p) = 1$, i.e., those beliefs give zero weight to level-1 dominated strategies. Moreover, the expectation within the integral is taken with respect to the common prior conditional on $\mathcal{I}_p$ which includes player p's type. Hence, exploiting this monotonicity, it is easy to see that for an outside observer, the subset of level-2 rationalizable strategies must satisfy

$$\mu_1 \in \Big[-\alpha_1 E_{\mathbb{P}_1}\big[\mathbb{1}\{t_2 \geq -\alpha_2\}|\mathcal{I}_1\big], -\alpha_1 E_{\mathbb{P}_1}\big[\mathbb{1}\{t_2 \geq 0\}|\mathcal{I}_1\big] = \Big[-\alpha_1\big(1 - \mathbb{P}_{t_1}(-\alpha_2)\big), -\alpha_1\big(1 - \mathbb{P}_{t_1}(0)\big)\Big]$$

and

$$\mu_2 \in \Big[-\alpha_2 E_{\mathbb{P}_2}\big[\mathbb{1}\{t_1 \geq -\alpha_1\}|\mathcal{I}_2\big], -\alpha_2 E_{\mathbb{P}_2}\big[\mathbb{1}\{t_1 \geq 0\}|\mathcal{I}_2\big] = \Big[-\alpha_2\big(1 - \mathbb{P}_{t_2}(-\alpha_1)\big), -\alpha_2\big(1 - \mathbb{P}_{t_2}(0)\big)\Big]$$

By induction, it is easy to prove the following claim.

Claim 1 *If $\alpha_p \leq 0$, a strategy of the type $Y_p = \mathbb{1}\{t_p \geq \mu_p\}$ is level-k rationalizable if and only if μ_1 and μ_2 satisfy*

$$\begin{aligned} &\mu_p \in \big[0, -\alpha_{-p}\big] \equiv \big[\mu^L_{p,1}, \mu^U_{p,1}\big], \quad \textit{for } k = 1 \textit{ and } p \in \{1,2\}. \\ &\mu_1 \in \Big[-\alpha_1 E\big[\mathbb{1}\{t_2 \geq \mu^U_{2,k-1}\}|\mathcal{I}_1\big], -\alpha_1 E\big[\mathbb{1}\{t_2 \geq \mu^L_{2,k-1}\}|\mathcal{I}_1\big]\Big] \equiv \big[\mu^L_{1,k}, \mu^U_{1,k}\big], \quad \textit{for } k > 1 \\ &\mu_2 \in \Big[-\alpha_2 E\big[\mathbb{1}\{t_1 \geq \mu^U_{1,k-1}\}|\mathcal{I}_2\big], -\alpha_2 E\big[\mathbb{1}\{t_1 \geq \mu^L_{1,k-1}\}|\mathcal{I}_2\big]\Big] \equiv \big[\mu^L_{2,k}, \mu^U_{2,k}\big], \quad \textit{for } k > 1 \end{aligned} \tag{4.6}$$

Remark 1 Any k-rational player is also k'-rational for any $1 \leq k' \leq k-1$. Also, for $p \in \{1,2\}$, with probability one, we have $\big[\mu^L_{p,k}, \mu^U_{p,k}\big] \subseteq \big[\mu^L_{p,k-1}, \mu^U_{p,k-1}\big]$ for any $k > 1$. Note also that these bounds are a function of I_p, the information player p conditions his beliefs on.

The two statements in Remark 1 follow because conditional on $\mathcal{I}_p$, the support $\mathbb{S}(\widehat{G}_p)$ of a k-level rational player is contained in that of a $k-1$-level rational player. In fact, if there is a unique BNE (conditional on $\mathcal{I}_p$), then $\mathbb{S}(\widehat{G}_p)$ would collapse to the singleton given by BNE beliefs as $k \to \infty$. Whenever it is warranted, we will clarify whether a k-level rational player is "at most k-level rational" or "at least k-level rational". For inference based on level-2 rationality, we can use inequalities similar to (4.3) above to map the observed choice probabilities to the predicted ones. In particular, we can use the thresholds from Claim 1 above to construct a map between the model and the observable outcomes using (4.4) above. This is illustrated in Figure 4.2 where we see that as one moves from level 1 to level 2, the middle square shrinks. In the next section, we parametrize the model to allow for observable heterogeneity and provide sufficient point identification conditions.

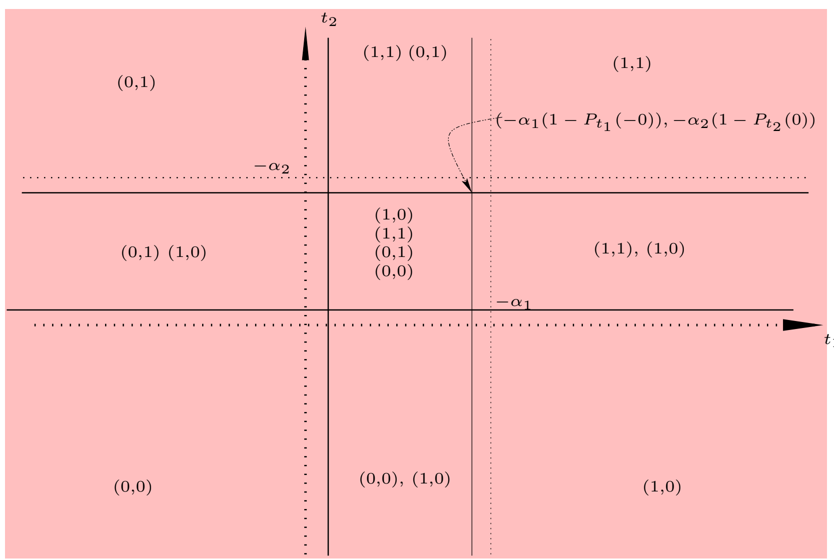


Figure 5: Observable Implications of Level-2 Rationality

4.3 Constructive Identification in a Parametric Model

From now on, we will express t_p as $t_p = X_p'\beta_p - \varepsilon_p$, where X_p is observable to the econometrician, ε_p is not, and β_p must be estimated. Throughout, we will assume $(\varepsilon_1, \varepsilon_2)$ to be continuously distributed, with scale normalized to one and a distribution of *known* functional form that depends on an unknown parameter that is part of the parameter vector of interest. For simplicity, we assume that ε_1 is independent of ε_2 where each has a CDF $H_p(.)$ for $p = 1, 2$. We will provide first an objective function that can be used to construct the identified set. This function will depend on the level k of rationality that the econometrician assumes ex-ante. We then discuss the identification of k. After that, we provide a set of sufficient conditions that will guarantee point identification under some assumptions. These point identification results provide insights into the kind of "variation" that is needed to shrink the identified set to a point.

As in the previous section, we make a common prior assumption. This assumption is only needed to compute bounds on beliefs for levels of rationality k that are strictly larger than 1. We will assume that player p observes ε_p and knows that ε_{-p} is a random variable with distribution H_{-p}. In principle, one can allow for correlated types, i.e., that H_{-p} is a function of p (this would accommodate strategic situations in which if one observes a "large" ε, then one thinks that his/her opponents' ε is also large.) But, we abstract from this for simplicity. We also implicitly assume that the econometrician knows the common prior distribution. Given this setup, we can iteratively construct bounds on the beliefs which will allow us to do inference.

Iterative construction of beliefs: We construct sets of consistent beliefs iteratively as follows. Let $\mathcal{I}$ be the information set for both players that we assume here is also observed by the econometrician. Typically, $\mathcal{I}$ contains the set of regressors X. Let $\pi^L_{-p}(\theta|k=1,\mathcal{I})=0$ and $\pi^U_{-p}(\theta|k=1,\mathcal{I})=1$. For $k>1$, let

$$\begin{aligned}\pi^L_1(\theta|k,\mathcal{I}) &= H_1\big(X_1'\beta_1+\alpha_1\pi^U_2(\theta|k\text{-}1,\mathcal{I})\big);\ \pi^U_1(\theta|k,\mathcal{I}) = H_1\big(X_1'\beta_1+\alpha_1\pi^L_2(\theta|k\text{-}1,\mathcal{I})\big)\\ \pi^L_2(\theta|k,\mathcal{I}) &= H_2\big(X_2'\beta_2+\alpha_2\pi^U_1(\theta|k\text{-}1,\mathcal{I})\big);\ \pi^U_2(\theta|k,\mathcal{I}) = H_2\big(X_2'\beta_2+\alpha_2\pi^L_1(\theta|k\text{-}1,\mathcal{I})\big)\end{aligned} \tag{4.7}$$

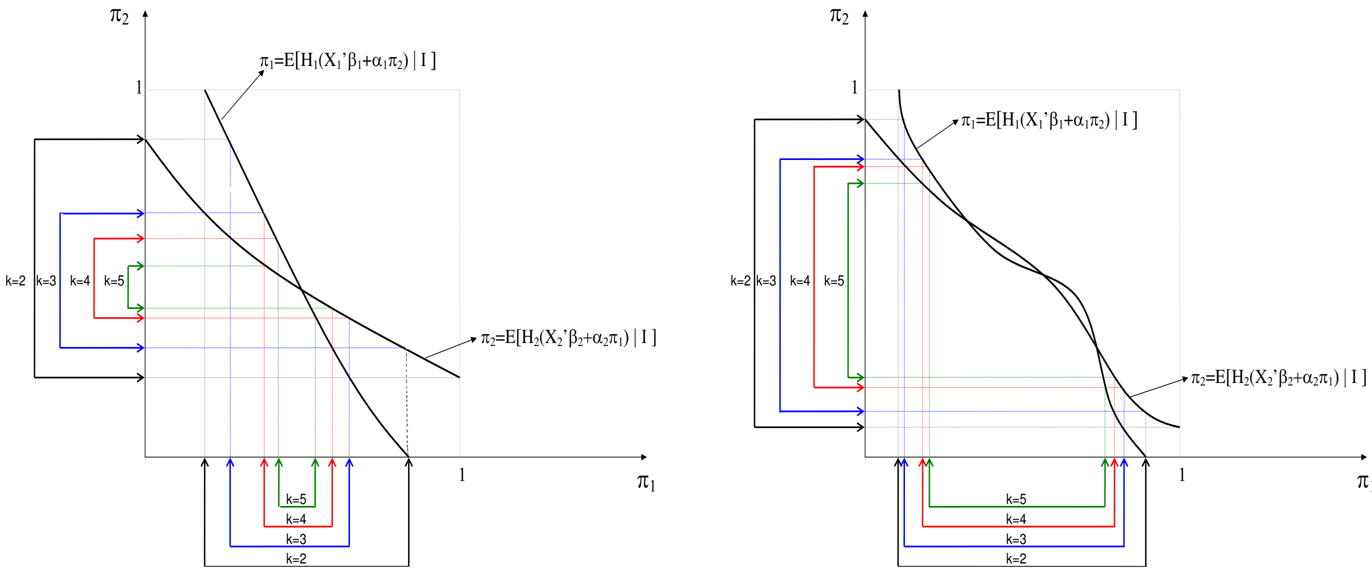
where $\pi^L_{-p}(\theta|k,\mathcal{I})$ and $\pi^U_{-p}(\theta|k,\mathcal{I})$ are the lower and upper bounds for level-k rationalizable beliefs by player p for $\Pr(Y_{-p}|\mathcal{I})$. In the case where we want to allow for correlation in types, then, the belief function for player p will depend on ε_p which would be part of a player specific information set and H_p would be the conditional CDF of $\varepsilon_p|\varepsilon_{-p}$. The econometrician does not observe either of the ε's. By induction, it is easy to show that

$$\big[\pi^L_{-p}(\theta|k;\mathcal{I}),\pi^U_{-p}(\theta|k;\mathcal{I})\big]\subseteq\big[\pi^L_{-p}(\theta|k\text{-}1;\mathcal{I}),\pi^U_{-p}(\theta|k\text{-}1;\mathcal{I})\big]\quad \text{w.p.1 in } \mathbb{S}(\mathcal{I}). \tag{4.8}$$

It also holds even if players condition on different information sets. Figure 6 depicts this case for a fixed realization I, a given parameter vector θ and $k\in\{1,2,3,4,5\}$. The left display in Figure 6 provides the belief iterations with a *unique* BNE while the right display shows the iterations with *multiple* BNE. Notice that when the game has a unique BNE, then as k

increases to infinity, the rationalizable beliefs will converge to the equilibrium ones. On the other hand, if the game has multiple BNE, then as $k \to \infty$, the beliefs converge to the outer equilibria as can be seen in the rhs display of Figure 6.

Figure 6: Rationalizable Beliefs for $k = 2, 3, 4$ and 5 :



Bounds for level-k rationalizable beliefs when $\mathcal{I}_1 = \mathcal{I}_2 \equiv \mathcal{I}$ (players condition on the same set of signals). Vertical axis shows level-k rationalizable bounds for player 1's beliefs about $\Pr(Y_2 = 1|\mathcal{I})$. Horizontal axis shows the equivalent objects for player 2. The graphs correspond to a particular realization $\mathcal{I}$ and a given parameter value θ.

Player p is level-k rational if and only if with probability one[2]

$$\mathbb{1}\big\{X_p'\beta_p + \alpha_p \pi_{-p}^U(\theta|k;\mathcal{I}) \geq \varepsilon_p\big\} \leq \mathbb{1}\big\{Y_p = 1\big\} \leq \mathbb{1}\big\{X_p'\beta_p + \alpha_p \pi_{-p}^L(\theta|k;\mathcal{I}) \geq \varepsilon_p\big\} \qquad (4.9)$$

This is is exactly the map that we can use to conduct inference. Let $W_p \equiv (X_p, \mathcal{I})$ and fix any k. For two vectors of scalars $a, b \in \mathbb{R}^{\dim(W_p)}$ let

$$\begin{aligned}
&\Lambda_p\big(\theta\big|a,b;k\big) = \\
&E\Bigg[\bigg(1 - \mathbb{1}\Big\{H_p\big(X_p'\beta_p + \alpha_p\pi_{-p}^U(\theta|k;\mathcal{I})\big) \leq \Pr(Y_p = 1|W_p) \leq H_p\big(X_p'\beta_p + \alpha_p\pi_{-p}^L(\theta|k;\mathcal{I})\big)\Big\}\bigg) \\
&\qquad \times \mathbb{1}\big\{a \leq W_p \leq b\big\}\Bigg]; \\
&\qquad\qquad \Gamma_p\big(\theta\big|k\big) = \int\int \Lambda_p\big(\theta\big|a,b;k\big)dF_{W_p}(a)dF_{W_p}(b); \quad \Gamma\big(\theta\big|k\big) = \Big(\Gamma_1\big(\theta\big|k\big)\ ,\ \Gamma_2\big(\theta\big|k\big)\Big)'
\end{aligned}$$

[2]Recall that we are studying the case $\alpha_p \leq 0$ for $p = 1, 2$.

$$(4.10)$$

where the inequality $a \leq W_p \leq b$ is element-wise and $W_p \sim F_{W_p}(\cdot)$. Take any conformable, positive definite matrix Ω and define

$$\Theta(k) = \Big\{\theta \in \underset{\theta}{\operatorname{argmin}}\ Q(\theta|k) = \underset{\theta}{\operatorname{argmin}}\ \Gamma\big(\theta\big|k\big)'\Omega\Gamma\big(\theta\big|k\big)\Big\} \tag{4.11}$$

By construction, $\Theta(k+1) \subseteq \Theta(k)$ for all k. If all we know is that all players are level-k rational, $\Theta(k)$ is the identified set. Methods meant for set inference can be used to construct a sample estimator of $\Theta(k)$ based on a random sample. Notice also that as compared with the Bayesian Nash solution, here one does not need to solve a fixed point map to obtain the equilibrium. Rather, rationalizability requires restrictions on player beliefs which can be implemented iteratively. We formally show below that $\Theta(k)$ contains the set of BNE for any $k > 0$.

Remark 2 Note that when $k = 1$, one does not need to specify the common prior assumption since beliefs here play no role. Hence, results will be robust to this assumption. However, depending on the magnitude of the α_p's, the bounds on choice probabilities predicted by such a model (where $k = 1$) can be wide.

Any player who is level-k rational is also level-k' rational for all $1 \leq k' \leq k-1$. As we pointed out above, an immediate consequence is that the identified sets satisfy $\Theta(k+1) \subseteq \Theta(k)$ for all k. A question of interest is how to do sharper inference based on level-k rationality by exploiting the features of the lower-level bounds. Specifically, if all players in the population are level-k rational, the structural parameter θ determines not only the features of the level-k bounds, but also how they relate to the features of the level-k' bounds for each $k' \leq k$. Exploiting this fact to generate a sharper identified set based on level-k rationality is a highly relevant question (see also Footnote 3).

On the rationality level k: Even though we do not explicitly study estimation and inference in this paper, the method above is meant to serve as the basis for a consistent estimate of the identified set $\Theta(k)$ for a given k. However, it is possible to learn something about k from a random sample. Specifically, suppose there exists k_0 such that players are at most level-k_0 rational. This describes a situation where the level-k_0 bounds hold with probability one, but the level-(k_0+1) bounds are violated with positive probability in the population. If we assume $k_0 \geq 1$ (the only interesting case), one can start with $k = 1$ and construct $\Theta(1)$ (as defined in 4.11). Next, for any $k \geq 2$ define

$$\underline{Q}(k) = \min_{\theta \in \Theta(1)} Q(\theta|k), \tag{4.12}$$

where $Q(\theta|k)$ is as defined in (4.11). Then,

(i) $\underline{Q}(k) = 0$ for all $k \leq k_0$. However, $\underline{Q}(k) = 0$ does not imply $k \leq k_0$.

(ii) $\underline{Q}(k) > 0$ implies $k > k_0$.

Suppose that different observations in the data set correspond to a game with a different level of rationality, then if $\underline{Q}(k) > 0$ and $\underline{Q}(k-1) = 0$, one would reject the hypothesis[3] that all the population is at least level k rational. If we assume ex-ante that $k_0 \geq \underline{k} > 1$ we could simply replace $\Theta(1)$ with $\Theta(\underline{k})$ in the definition of $\underline{Q}(k)$ in Equation (4.12). Alternatively, in settings where at least a subset of the structural parameter θ is known (e.g, experiments), we could evaluate if players are at least level-k_0 rational by testing whether or not $\theta_0 \in \Theta(k_0)$ (the identified set for level-k_0 rationality). Otherwise, a test that would fail to reject $\Theta(k_0+1) = \emptyset$ would indicate that players are at most level-k_0 rational.

4.4 Sufficient point identification conditions

In this section, we study the problem of point identification of the parameter of interests in the game above. In particular, we provide sufficient point identification conditions for level-1 rational play and for levels $k > 1$. These conditions can provide insights about what is required to shrink the identified set to a point (or a vector). Here, we allow for the information sets to be different, i.e., that player p conditions on $\mathcal{I}_p$ when making decisions and allow for exclusion restrictions where $\mathcal{I}_1 \neq \mathcal{I}_2$. We start with sufficient conditions for level 1 rationalizability.

4.4.1 Identification with level-1 Rationalizability

Let $\theta_p = (\beta_p, \alpha_p)$ and $\theta = (\theta_1, \theta_2)$, we have the following identification result.

Theorem 1 *Suppose X_p has full rank for $p = 1, 2$ and let $X \equiv (X_1, X_2)$, assume $\alpha_p < 0$ for $p = 1, 2$ and let Θ denote the parameter space. Let there be a random sample of size N from the game above. Consider the following condition.*

[3]Strictly speaking, this would be a joint test of the rationality hypothesis and all other maintained assumptions.

A1.i For each player p, there exists a continuously distributed $X_{\ell,p} \in X_p$ with nonzero coefficient $\beta_{\ell,p}$ and unbounded support conditional on $X \setminus X_{\ell,p}$ such that for any $c \in (0,1)$, $b \neq 0$ and $q \in \mathbb{R}^{dim(X_{-\ell,p})}$, there exists $C_{b,q,m} > 0$ such that

$$Pr(\varepsilon_p \leq bX_{\ell,p} + q'X_{-\ell,p}|X) > m \;\forall\; X_{\ell,p} : \; sign(b) \cdot X_{\ell,p} > C_{b,q,m}. \tag{4.13}$$

A.1-ii For $p = 1, 2$, let $X_{d,p}$ denote the regressors that have bounded support but are not constant. Suppose Θ is such that for any $\beta_{d,p}, \widetilde{\beta}_{d,p} \in \Theta$ with $\widetilde{\beta}_{d,p} \neq \beta_{d,p}$ and for any $\alpha_p \in \Theta$,

$$Pr\Big(\big|X'_{d,p}(\beta_{d,p} - \widetilde{\beta}_{d,p})\big| > |\alpha_p| \;\Big|X \setminus X_{d,p}\Big) > 0. \tag{4.14}$$

If all we know is that players are Level-1 rational :

(a) If (A1.i) holds, the coefficients $\beta_{\ell,p}$ are identified.

(b) If (A1.ii) holds, the coefficients $\beta_{d,p}$ are identified.

(c) We say that player p is pessimistic with positive probability if for any $\Delta > 0$, there exists $\mathcal{X}_\Delta \in \mathbb{S}(X_p)$ such that $Pr(Y_p = 1|X) < Pr(\varepsilon_p \leq X_p'\beta_{b_0} + \alpha_{p0}|X) + \Delta$ whenever $X_p \in \mathcal{X}_\Delta$. If (A1.i-ii) holds and player p is pessimistic with positive probability, the identified set[4] for α_p is $\{\alpha_p \in \Theta : \alpha_p \leq \alpha_{p0}\}$.

The results in Theorem 1 imposed no restrictions on $\mathcal{I}_p$. In particular, players can condition their beliefs on unobservables (to the econometrician). A special case of condition (A1.i) is when ε_p is independent of X. The condition in (A1.ii) says how rich the support of the bounded shifters must be in relation to the parameter space. Covariates with unbounded support satisfy this condition immediately given the full-rank assumption. Finally, similar identification results to Proposition 1 hold for cases $\alpha_p \geq 0$ and $\alpha_1\alpha_2 \leq 0$. The proof of the above theorem is given in an appendix.

4.4.2 Identification with "$k \geq 2$-level" rationalizability

We now move on to the case of rationalizable beliefs of higher order. Our goal is to investigate if a higher degree of rationality will help in the task of point identifying α_p. To simplify the analysis, we will assume from now on that ε_p is independent of X and of $\mathcal{I} \equiv (\mathcal{I}_1, \mathcal{I}_2)$. This

[4] Here, we refer to the identified set as the set of values of α_p that are observationally equivalent, conditional on observables, to the true value α_{p_0}.

assumption could be replaced with one along the lines of (A1) in theorem 1. We make the assumption that $\mathcal{I}$ is observed by the econometrician. We will relax this assumption in a later section. Again, let the common prior assumption be denoted by $H_p(.)$. The beliefs of the players for any level k rationality can be constructed as we did in the previous section. Our point identification sufficient conditions are summarized in theorem 2 below.

Theorem 2 *Suppose there exists a subset $\mathcal{X}_1^* \subseteq \mathbb{S}(X_1)$ where X_1 has full-column rank such that for any $X_1 \in \mathcal{X}_1^*$, $\varepsilon > 0$ and $\theta_2 \in \Theta$, there exist $\Im_{1_\varepsilon}^* \subset \mathbb{S}(\mathcal{I}_1|X_1)$ and $\Im_{1_\varepsilon}^{**} \subset \mathbb{S}(\mathcal{I}_1|X_1)$ such that*

$$
\begin{aligned}
&Max\Big\{1 - E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big] - E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\Big\} < \varepsilon \quad \forall\, \mathcal{I}_1 \in \Im_{1_\varepsilon}^* \\
&Max\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big] - E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\Big\} < \varepsilon \quad \forall\, \mathcal{I}_1 \in \Im_{1_\varepsilon}^{**}
\end{aligned}
\tag{4.15}
$$

A special case in which (4.15) holds is when there exists $X_{2_\ell} \in (X_2 \cap W_1)$ with nonzero coefficient in Θ such that X_{2_ℓ} has unbounded support conditional on $(X_2 \cup W_1) \setminus X_{2_\ell}$. We could refer to (4.15) as an "informative signal" condition. Note that implicit in (4.15) is an exclusion restriction in the parameter space that precludes $\beta_2 = 0$ for any $\theta_2 \in \Theta$. If (4.15) holds, then for any $\theta \in \Theta$ such that $\theta_1 \neq \theta_{1_0}$, there exists either $\mathcal{W}_1^* \subset \mathbb{S}(W_1)$ or $\mathcal{W}_1^{**} \subset \mathbb{S}(W_1)$ such that

$$
\begin{aligned}
&H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k;\mathcal{I}_1)\big) < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^U(\theta_0|k;\mathcal{I}_1)\big) \quad \forall\, W_1 \in \mathcal{W}_1^* \text{ and all } k \geq 2, \\
&H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k;\mathcal{I}_1)\big) > H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k;\mathcal{I}_1)\big) \quad \forall\, W_1 \in \mathcal{W}_1^{**} \text{ and all } k \geq 2.
\end{aligned}
\tag{4.16}
$$

Therefore, for any $k \geq 2$ the Level-k rationalizable bounds for Player 1's conditional choice probability of $Y_1 = 1|W_1$ that correspond to θ will be disjoint with those of θ_0 with positive probability. As a consequence, if (4.15) holds and the population of Players 1 are at least Level-2 rational, θ_{1_0} is identified. By symmetry, θ_{2_0} will be point-identified if the above conditions hold with the subscripts "1" and "2" interchanged.

For the case in which $\mathcal{I}_1 = \mathcal{I}_2 = X$, Figures 7 and 8 illustrate four graphical examples of how the "informative signals" condition (4.15) in Theorem 2 yields disjoint Level-2 bounds.

The ability to shift the upper and lower bounds for Level-2 rationalizable beliefs arbitrarily close to 1 or 0 is essential for the point-identification result in Theorem 2. For

Figure 7: Graphical Examples of Informative Signals I

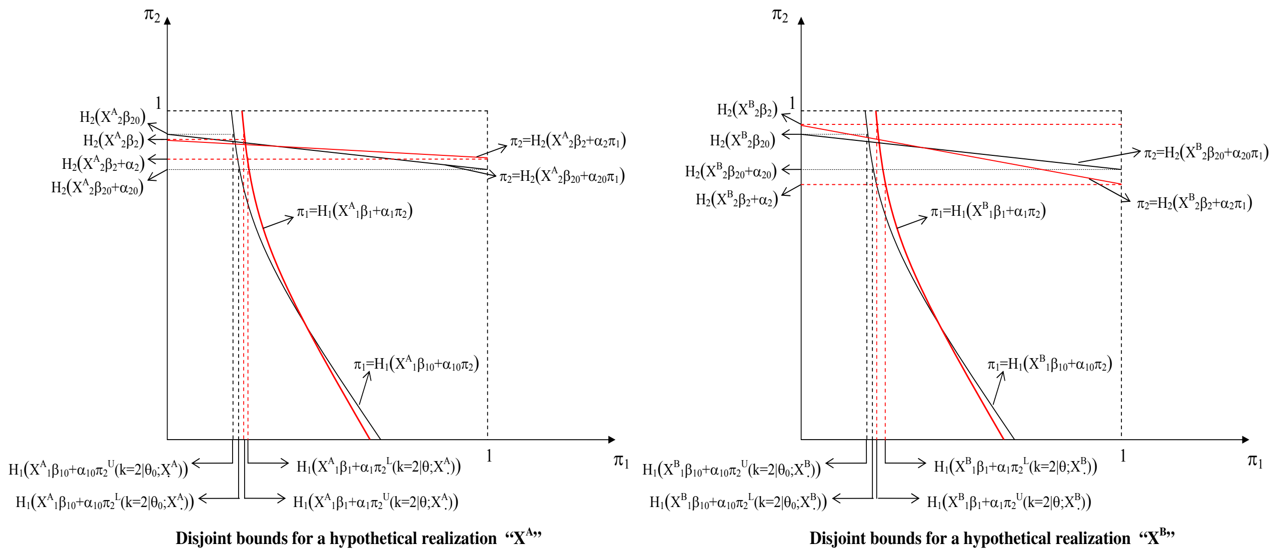
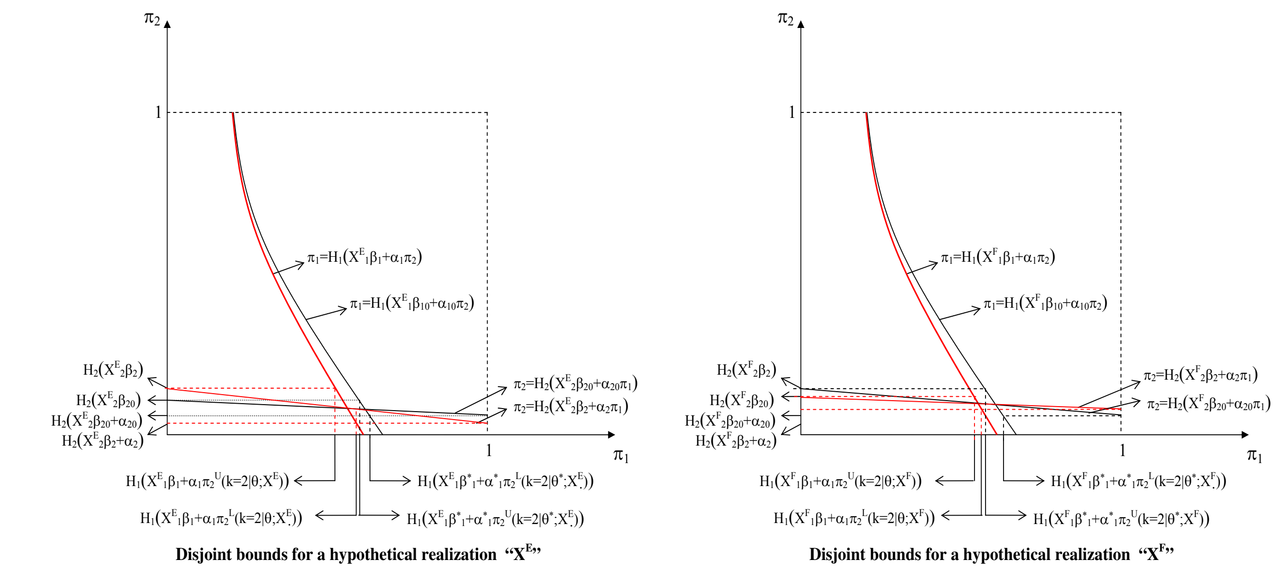


Figure 8: Graphical Examples of Informative Signals II



simplicity, the intercept Δ_1 is subsumed in $X_1'\beta_1$ in the labels of these figures.

Inference on the rationality level k with point-identification of $\boldsymbol{\theta}$: If θ is point-identified, the upper rationality bound k_0 defined prior to Equation (4.12) is also point-identified. This follows because $Q(\theta_0|k) = 0$ if and only if $k \leq k_0$, where $Q(\theta|k)$ is defined in Equation (4.11). To see why this is not true when θ is set-identified, go back to parts (i)-(ii) following Equation (4.12).

4.5 Bayesian-Nash Equilibria and Rationalizable Beliefs

As before, let $\mathcal{I}_p$ be the signal player p uses to condition his beliefs about his opponent's expected choice, and let $\mathcal{I} \equiv (\mathcal{I}_1, \mathcal{I}_2)$. The set of Bayesian-Nash equilibria (BNE) is defined as any pair $\big(\pi_1^*(\mathcal{I}_2), \pi_2^*(\mathcal{I}_1)\big) \equiv \pi^*(\mathcal{I})$ that satisfies

$$\begin{aligned} \pi_1^*(\mathcal{I}_2) &= E\big[H_1(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1))\big|\mathcal{I}_2\big] \\ \pi_2^*(\mathcal{I}_1) &= E\big[H_2(X_2'\beta_2 + \alpha_2\pi_1^*(\mathcal{I}_2))\big|\mathcal{I}_1\big] \end{aligned} \tag{4.17}$$

By construction, the set of rationalizable beliefs for $\mathcal{I}$ must include the BNE set for any rational level k. The following result formalizes this claim.

Proposition 1 *Let*

$$\mathcal{R}(\mathcal{I};k) = \Big[\pi_1^L(\theta|k;\mathcal{I}_2), \pi_1^U(\theta|k;\mathcal{I}_2)\Big] \times \Big[\pi_2^L(\theta|k;\mathcal{I}_1), \pi_2^U(\theta|k;\mathcal{I}_1)\Big]$$

denote the set of level-k rationalizable beliefs. Then, with probability one, the BNE set described in (4.17) *is contained in* $\mathcal{R}(\mathcal{I};k)$ *for* any $k \geq 1$.

We present the proof for the case $\alpha_p \leq 0$ for $p = 1, 2$, which we have focused on. The proof can be adapted to all other cases. We will proceed by induction by proving first the following claim.

Claim 2 Let $\pi^*(\mathcal{I}) \equiv \big(\pi_1^*(\mathcal{I}_2)\ ,\ \pi_2^*(\mathcal{I}_1)\big)$ be any BNE. Then, for any $k \geq 1$, with probability one, we have: $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k)$ implies $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k+1)$ w.p.1.

Proof of Claim 2: If $\alpha_p = 0$ for $p = 1$ or $p = 2$, the result follows trivially. Suppose $\alpha_1 = 0$, then $\pi_1^L(\theta|k;\mathcal{I}_2) = \pi_1^U(\theta|k;\mathcal{I}_2) = \pi_1^*(\mathcal{I}_2) = E\big[H_1(X_1'\beta_1)\big|\mathcal{I}_2\big]$ and $\pi_2^L(\theta|k;\mathcal{I}_1) = \pi_2^U(\theta|k;\mathcal{I}_1) = \pi_2^*(\mathcal{I}_1) = E\big[H_2(X_2'\beta_2 + \alpha_2\pi_1^*(\mathcal{I}_2))\big|\mathcal{I}_1\big]$ for all $k \geq 1$. We focus on the case $\alpha_p < 0$ for $p = 1, 2$. Now, suppose $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k)$, but $\pi^*(\mathcal{I}) \notin \mathcal{R}(\mathcal{I};k)$. Suppose for example that $\pi_1^L(\theta|k+1;\mathcal{I}_2) > \pi_1^*(\mathcal{I}_2)$. Since $\alpha_1 < 0$, this can be true if and only if

$$\underbrace{E\Big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^U(\theta|k;\mathcal{I}_1)\big)\Big|\mathcal{I}_2\Big]}_{=\pi_1^L(\theta|k+1;\mathcal{I}_2)} > \underbrace{E\Big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1)\big)\Big|\mathcal{I}_2\Big]}_{=\pi_1^*(\mathcal{I}_2)}$$

For this inequality to be satisfied, it *cannot* be the case that $\pi_2^*(\mathcal{I}_1) \leq \pi_2^U(\theta|k;\mathcal{I}_1)$. But this violates the assumption that $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k)$. Therefore, we must have $\pi_1^L(\theta|k+1;\mathcal{I}_2) \leq$

$\pi_1^*(\mathcal{I}_2)$. Suppose now that $\pi_1^U(\theta|k;\mathcal{I}_2) < \pi_1^*(\mathcal{I}_2)$. This can be true if and only if

$$\underbrace{E\Big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^L(\theta|k;\mathcal{I}_1)\big)\Big|\mathcal{I}_2\Big]}_{=\pi_1^U(\theta|k+1;\mathcal{I}_2)} < E\Big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1)\big)\Big|\mathcal{I}_2\Big]$$

For this inequality to be satisfied, it *cannot* be the case that $\pi_2^*(\mathcal{I}_1) \geq \pi_2^L(\theta|k;\mathcal{I}_1)$. Once again, this violates the assumption $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k)$. Therefore, we must have $\pi_1^U(\theta|k+1;\mathcal{I}_2) \geq \pi_1^*(\mathcal{I}_2)$. These results imply that we must have $\pi_1^L(\theta|k+1;\mathcal{I}_2) \leq \pi_1^*(\mathcal{I}_2) \leq \pi_1^U(\theta|k+1;\mathcal{I}_2)$. Following the same steps we can establish that we must have $\pi_2^L(\theta|k+1;\mathcal{I}_1) \leq \pi_2^*(\mathcal{I}_1) \leq \pi_2^U(\theta|k+1;\mathcal{I}_1)$. Combined, these yield $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k+1)$ as claimed. $\square$

Proof of Proposition 1: Follows from Claim 2 and the fact that level-1 rational players satisfy $H_p(X_p'\beta_p + \alpha_p) \leq E[Y_p|X_p] \leq H_p(X_p'\beta_p)$, which yields $\mathcal{R}(\mathcal{I};k=1) = [0,1]\times[0,1]$. Consequently $\mathcal{R}(\mathcal{I};k=1)$ contains all BNE. It follows from Claim 2 that $\mathcal{R}(\mathcal{I};k=1)$ contains all BNE for all $k \geq 1$. $\square$

BNE vs Rationalizability: Identification: Naturally, it is always guaranteed that one gets a weakly *smaller* identified set with BNE assumptions since the predicted outcomes based on equilibrium use stronger assumptions on player beliefs. The "size" of the rationalizable outcome set depends on the distance between the "smallest" and the "largest" equilibrium. In the case of a unique equilibrium, one can see that in the above game and as $k \to \infty$, the predicted outcomes under both solution concepts converge. In addition, in the simple example above, predicted outcomes based on rationality of order k, for any k, are a lot easier to solve for since they do not require solutions to fixed point problems especially in cases of multiple equilibria.

5 Identification in First Price IPV Auctions with Rationalizable Bids

This section considers a situation in which a population of symmetric, risk-neutral potential buyers must bid simultaneously for a single good. We focus on a first-price auction with independent private values, although our results can be adapted to the case of interdependent private values and affiliated signals. As it is usually the case in the econometric analysis of auctions, the object of interest is the distribution of private values. Under the assumption that observed bids conform to a Bayesian-Nash equilibrium (BNE), nonparametric point identification for this distribution has been established for example by Guerre, Perrigne, and Vuong (1999). Hence, equilibrium assumptions (and other conditions) deliver point identification of the valuation distribution. Here, we relax the BNE requirement and assume only that buyers are strategically sophisticated in the sense of Battigalli and Siniscalchi (2003), abbreviated henceforth as BS. Other strategic assumptions that can be used and that deliver qualitatively different results than BS's interim rationalizability is the $\mathcal{P}$ dominance concept introduced for auctions setups by Dekel and Wolinsky (2003) and more recently Crawford and Iriberri (2007b). Here, we just highlight what can be learned with the BS setup and compare those to Bayesian Nash equilibrium. Again, we do not advocate interim rationalizability as the basis for strategic play in auctions. We just consider one other strategic environment and study its implication on learning about the valuation distribution from observed data on bids.

BNE requires rational, expected utility maximizing buyers with correct beliefs. Strategically sophisticated buyers are rational and expected utility maximizers, but their beliefs may or may not be correct. This characterization includes BNE as a special case. The degree of sophistication will be characterized using the concept of interim rationalizability. As we will see, this will lead to the notion of "Level-k rationalizable bids" for $k \in \mathbb{N}$. We describe these concepts next.

Let $F_0(\cdot)$ denote the distribution of v_i, the private valuation of bidder i. We assume $F_0(\cdot)$ to be common knowledge among the bidders, and focus on the case where $F_0(\cdot)$ is log-concave and absolutely continuous with respect to Lebesgue measure. We assume its support to be of the form $[0, \omega)$ (i.e, normalize its lower bound by zero) and allow, in principle, the case $\omega = +\infty$. Assume for the moment that the seller's reservation price p_0 is equal to zero. We will explicitly introduce a strictly positive reservation price below.

Assumptions about Bidders' Beliefs We make the following assumptions concerning bidders' beliefs.

Following BS, we assume that bidders expect all positive bids to win with strictly positive probability and this is common knowledge. This condition will ensure that it is common knowledge that *no bidder will bid beyond his/her valuation irrespective of his beliefs.* It also implies that every bidder with nonzero private value will submit a strictly positive bid[5].

[5] Interim rationalizability will only naturally produce upper bounds for rationalizable bids. Additional, ad-hoc assumptions could be made to characterize a lower bound.

Therefore, with probability one the number of potential bidders $\mathcal{N}$ is equal to the number of actual bidders (only a bidder with valuation equal to zero is indifferent between entering the bid or not). We restrict further attention to beliefs that assign positive probability only to increasing bidding functions. Formally, let $\mathcal{B}$ denote the space of all functions of the form

$$\mathcal{B} = \big\{b(.) : [0,\omega) \to \mathbb{R}_+ : \; b(v) \le v, \text{ and } v > v' \; \Rightarrow \; b(v) > b(v')\big\}. \tag{5.18}$$

We will let $\mathcal{N}$ denote the number of potential bidders in the population and denote $\mathcal{B}_{-i} = \mathcal{B}^{\mathcal{N}-1}$. Beliefs for bidder i are probability distributions defined over a sigma-algebra $\Delta_{\mathcal{B}_{-i}}$, where this sigma-algebra is such that singletons in $\mathcal{B}_{-i}$ are measurable[6]. A conjecture by bidder i is a degenerate belief that assigns probability mass one to a singleton $\{b_j\}_{j\neq i} \in \mathcal{B}_{-i}$. The distribution of valuations $F_0(\cdot)$ as well as $\mathcal{N}$ are common knowledge among potential bidders. This is similar to the common prior assumption made in the previous section.

As we will see, restricting attention to beliefs in $\mathcal{B}$ will yield rationalizable upper bounds for bids which also belong in $\mathcal{B}$. It also simplifies the analysis, for example, by ruling out ties in the characterization of players' expected utility. Finally, as we will argue below (and is formally shown in BS), restricting attention to beliefs in $\mathcal{B}$ will imply that Bayesian-Nash equilibrium (BNE) optimal bids are always rationalizable.

5.1 Prediction with Level-k Rationalizable bids

Here, we follow the setup in BS. We have a population of $\mathcal{N}$ risk-neutral potential buyers, bidding simultaneously for a single object. With a zero reservation price, we can interpret $\mathcal{N}$ as the number of observed bids that is common knowledge among the bidders. Each bidder i observes his valuation v_i, independent of those of other bidders with identical log-concave, continuous distribution $F_0(\cdot)$. The highest bid wins the object, ties are broken at random and only the winner pays his bid. The space of beliefs we focus on assigns probability zero to ties. Therefore, the decision problem for bidder i can be expressed as[7]

$$\max_{b\ge 0} \big(v_i - b\big)\widehat{\Pr}_i\Big[\max_{j\neq i} b(v_j) \le b\Big], \tag{5.19}$$

where $\widehat{\Pr}_i(\cdot)$ denotes bidder i's subjective probability, derived from his beliefs and knowledge of $F_0(\cdot)$. For level-1 rational bidding, any bidder i whose bids satisfy

$$b \le v_i \equiv \overline{B}_1(v_i;\mathcal{N}) \quad \text{w.p.1} \tag{5.20}$$

[6] The results we analyze here do not depend on the specific choice of the sigma algebra, as long as it satisfies the singleton-measurability mentioned here. See footnote 10 in BS.

[7] Strictly speaking, what matters is that ties have probability zero for the most pessimistic conjecture.

are called *level-1 rational bidders*. Any expected-utility maximizer bidder i must be level-1 rational regardless of whether or not his beliefs live in $\mathcal{B}_{-i}$. Hence we have

Result: (BS) Any bid with $b_i \leq v_i$ is level-1 rational

This was proved in BS where they also show that the bound is sharp, i.e., that for any bid in the bound, there exists a consistent and valid level 1 belief function for which that bid is a best response. This result is interesting since in this setup, one cannot bound the bids from below. This is in marked contrast to the BNE prediction. Note that the bound above depends on the continuity of the valuation and the assumption that any positive bid has a positive chance of winning. In another case where the valuations are assumed to take countable values, Dekel and Wolinsky (2003) showed that a form of rationalizability implies tight bounds on the bidding function in the limit as the number of bidders increases. Here, we will derive strategies for identification of $F(.)$ based on the BS results, but these strategies can be easily adapted to other strategic setups like ones suggested by Dekel and Wolinsky.

Higher order rationality: We now characterize the identified features in an auction with higher rationality levels. Focus on bidders with beliefs in $\mathcal{B}_{-i}$. The **most pessimistic assessment** in $\mathcal{B}_{-i}$ is given by the conjecture $b(v_j) = \overline{B}_1(v_j; \mathcal{N}) = v_j$ for all $j \neq i$ (the upper bound for bids for level-1 rational bidders). Since bidder i knows F_0, his optimal expected utility for this assessment is

$$\begin{aligned}\max_{b\geq 0} \big(v_i - b\big)\Pr\Big[\max_{j\neq i} \overline{B}_1(v_j;\mathcal{N}) \leq b\Big] &= \max_{b\geq 0}\big(v_i - b\big)\Pr\Big[\max_{j\neq i} v_j \leq b\Big] \\ &= \max_{b\geq 0} \big(v_i - b\big) F_0(b)^{\mathcal{N}-1} \qquad (5.21) \\ &\equiv \underline{\pi}_2^*(v_i;\mathcal{N})\end{aligned}$$

where $\underline{\pi}_2^*(v_i;\mathcal{N})$ is the lower bound for optimal expected utility (5.19) for all beliefs in $\mathcal{B}_{-i}$. The upper expected utility bound for an arbitrary bid b is trivially given by $(v_i - b)$ for any possible beliefs (no bidder would ever expect to win the good with probability higher than one). Any bid submitted by a rational (i.e, expected-utility maximizer) bidder with beliefs in $\mathcal{B}_{-i}$ must satisfy

$$v_i - b \geq \underline{\pi}_2^*(v_i;\mathcal{N}) \;\Rightarrow\; b \leq v_i - \underline{\pi}_2^*(v_i;\mathcal{N}) \equiv \overline{B}_2(v_i;\mathcal{N}) \quad \text{w.p.1} \qquad (5.22)$$

We refer to bidders who satisfy (5.22) as *level-2 rational bidders*. Given our assumptions, $\overline{B}_2(v_i;\mathcal{N})$ is increasing, concave and satisfies $\overline{B}_2(v_i;\mathcal{N}) \leq \overline{B}_1(v_i;\mathcal{N}) = v_i$, with strict inequality for all $v_i > 0$.[8] Therefore, $\overline{B}_2 \in \mathcal{B}$. Let $\overline{S}_2(\cdot;\mathcal{N})$ denote the inverse of $\overline{B}_2(\cdot;\mathcal{N})$.

[8]These and more properties are enumerated in BS, who focus on a more general case which allows for

We refer to *level-3 rational bidders* as those whose beliefs incorporate the level-2 upper bound (5.22). The most pessimistic assessment for level-3 rational bidders is the conjecture $b(v_j) = \overline{B}_2(v_j;\mathcal{N})$ for all $j \neq i$. The optimal expected utility for this pessimistic assessment is

$$\max_{b\geq 0}\,\big(v_i - b\big)\Pr\Big[\max_{j\neq i}\overline{B}_2(v_j;\mathcal{N})\leq b\Big] = \max_{b\geq 0}\,\big(v_i-b\big)F_0\big(\overline{S}_2(b;\mathcal{N})\big)^{\mathcal{N}-1}\equiv \underline{\pi}_3^*(v_i;\mathcal{N}) \quad (5.23)$$

Using the same logic that led to (5.22), the set of rationalizable bids for level-3 rational bidders must satisfy

$$v_i - b \geq \underline{\pi}_3^*(v_i;\mathcal{N}) \;\Rightarrow\; b\leq v_i - \underline{\pi}_3^*(v_i;\mathcal{N})\equiv \overline{B}_3(v_i;\mathcal{N}) \quad \text{w.p.1.} \tag{5.24}$$

The level-3 upper bound for rationalizable bids, $\overline{B}_3(\cdot;\mathcal{N})$ is increasing, concave and satisfies $\overline{B}_3(\cdot;\mathcal{N})\leq \overline{B}_2(\cdot;\mathcal{N})$, with strict inequality for nonzero valuations. To see why the last result holds, recall that $\overline{B}_2(v_i;\mathcal{N}) = v_i - \underline{\pi}_2^*(v_i;\mathcal{N})\equiv \overline{B}_1(v_i;\mathcal{N}) - \underline{\pi}_2^*(v_i;\mathcal{N})$. Therefore, for any b we have $\Pr\Big[\max_{j\neq i}\overline{B}_2(v_j;\mathcal{N})\leq b\Big]\geq \Pr\Big[\max_{j\neq i}\overline{B}_1(v_j;\mathcal{N})\leq b\Big]$. Immediately, this implies $\underline{\pi}_3^*(\cdot;\mathcal{N})\geq \underline{\pi}_2^*(\cdot;\mathcal{N})$ and therefore $\overline{B}_3(\cdot;\mathcal{N})\leq \overline{B}_2(\cdot;\mathcal{N})$. Since $F_0(\cdot)$ is not assumed to have point masses, all the above inequalities are strict for any $v_i>0$. Proceeding iteratively, the ***level-k bound for rationalizable bids*** is given by

$$\begin{aligned} &b_i\leq v_i - \underline{\pi}_k^*(v_i;\mathcal{N})\equiv \overline{B}_k(v_i;\mathcal{N}) \quad \text{w.p.1.,} \quad \text{where}\\ &\underline{\pi}_k^*(v_i;\mathcal{N}) = \max_{b\geq 0}\,\big(v_i-b\big)F_0\big(\overline{S}_{k-1}(b;\mathcal{N})\big)^{\mathcal{N}-1}, \end{aligned} \tag{5.25}$$

and $\overline{S}_{k-1}(\cdot;\mathcal{N})$ is the inverse function of $\overline{B}_{k-1}(\cdot;\mathcal{N})$. The level-k upper bounds for rationalizable bids, $\overline{B}_k(\cdot;\mathcal{N})$, are increasing, concave and satisfy $\overline{B}_{k+1}(v;\mathcal{N})\leq \overline{B}_k(v;\mathcal{N})$ for all k, with strict inequality for all $v>0$. Let $b^{BNE}(v;\mathcal{N})$ denote the optimal BNE bidding function, produced by self-consistent, correct beliefs. BS have shown that $\overline{B}_k(\cdot;\mathcal{N})\geq b^{\text{BNE}}(\cdot;\mathcal{N})$ for all $k\in\mathbb{N}$. In particular, this is true for $\lim_{k\to\infty}\overline{B}_k(\cdot;\mathcal{N})$, which is well-defined by the aforementioned monotonicity property of the sequence $\big\{\overline{B}_k(\cdot;\mathcal{N})\big\}_{k\in\mathbb{N}}$. Bidding below $b^{\text{BNE}}(\cdot;\mathcal{N})$ is always rationalizable for any rationality level k. All results presented here will be consistent with this type of behavior.

Example.- Suppose private values are exponentially distributed, with $F_0(v) = 1-\exp\{-\theta v\}$ and $\theta>0$. We have $F_0(v)/f_0(v) = \frac{1-\exp\{-\theta v\}}{\theta\exp\{-\theta v\}} = \frac{1}{\theta}\exp\{\theta v\} - \frac{1}{\theta}$, which is an increasing function of v for all $\theta>0$, establishing log-concavity of F_0. Figure 9 depicts $\overline{B}_k(\cdot;\mathcal{N})$, the level-k

interdependent values.

rationalizable bounds for bids for the case $\theta = -0.25$, $\mathcal{N} = 2$ (two bidders) and $k = 1, 2, 3, 4$. This graphical example illustrate the features described above for these bounds. Namely, $\overline{B}_k(\cdot; \mathcal{N})$, is continuous, increasing, concave, invertible and satisfies $\overline{B}_{k+1}(v; \mathcal{N}) \leq \overline{B}_k(v; \mathcal{N})$ for all k, with strict inequality for all $v > 0$. For this particular example, the bounds corresponding to $k \geq 5$ are graphically indistinguishable from $\overline{B}_4(v; \mathcal{N})$.

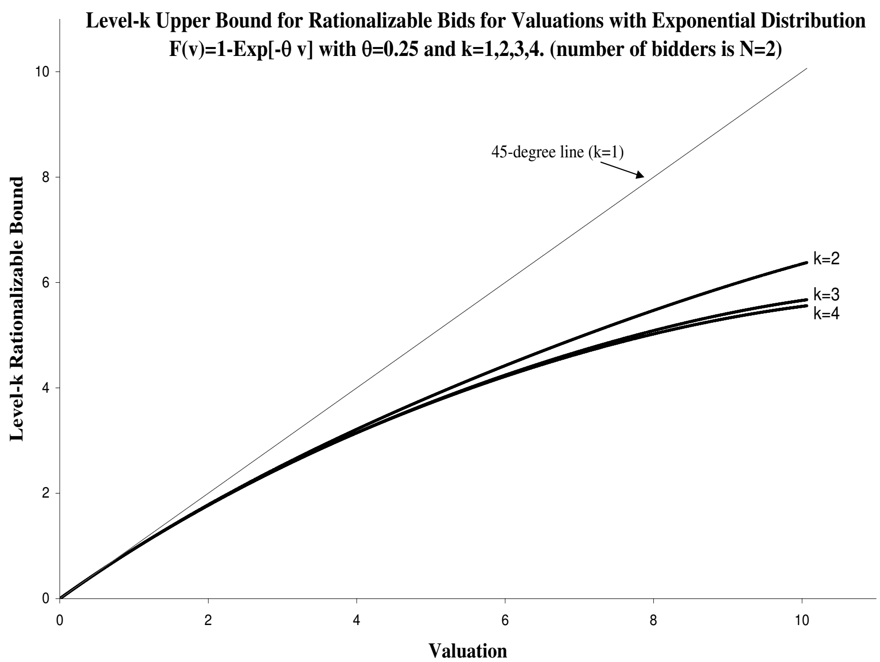


Figure 9: Level-k rationalizable bounds $\overline{B}_k(\cdot; \mathcal{N})$ for $F_0(v) = 1 - e^{-0.25v}$, $\mathcal{N} = 2$ and $k = 1, 2, 3, 4$.

5.2 Identification with level-k rationality

This section exploits the above bounds to learn about the distribution of valuation given a random sample of bids. Initially, we will assume that we observe a random sample of size L auctions each with $\mathcal{N}$ bidders with non binding reservation prices. An observation is an auction and for each auction, we assume also that we observe all the submitted bids. In the next sections we will relax the binding reserve price assumptions and we allow that only winning bids are observed.

We assume a semiparametric setting where F_0 belongs to a space of log-concave, absolutely continuous distribution functions with support $[0, \omega)$ of the form

$$\mathcal{F}_v^{\Theta} = \Big\{ F(\cdot;\theta) : \ \theta \in \Theta, \text{ and } F_0(\cdot) = F(\cdot;\theta_0) \text{ for some } \theta_0 \in \Theta \Big\} \tag{5.26}$$

Here, one can also think of Θ as a set of functions and hence the above definition accommodates nonparametric analysis. Denote the level-k upper bound that corresponds to $F(\cdot;\theta)$ by $\overline{B}_k(\cdot;\mathcal{N}|\theta)$.

Level-1 rationality: For rationality of level 1, the game predicts that

$$0 \le b_i^l \le v_i^l \text{ for all } i = 1, \ldots, \mathcal{N} \quad l = 1, \ldots, L$$

This is a problem of inference with interval data. The b's are observed and the v's are not, but we observe a bound on every observation. The object of interest is the distribution function F of the valuations v (here, one can introduce auction heterogeneity that is observed). This implies that

$$F_0(t;\theta) \equiv P(v \le t) \le P(b \le t) \equiv G_b(t)$$

So, with the first level of rationality, we can bound the valuation distribution above by the observed distribution of the bids. Inference here will be handled below and is based on replacing the observed bids distribution with its consistent empirical analog.

Level-k rationality: Similarly to above, for level-k and any $\theta \in \Theta$, we have

$$\begin{aligned} &0 \le b_i^l \le v_i^l - \underline{\pi}_k^*(v_i;\mathcal{N}|\theta) \equiv \overline{B}_k(v_i;\mathcal{N}) \\ &\quad \text{for all } i = 1, \ldots, \mathcal{N} \quad l = 1, \ldots, L \end{aligned}$$

Hence, this means that if bidders are level-k rational,

$$F\big(\overline{S}_k(t;\mathcal{N}|\theta_0);\theta_0\big) \le P(b \le t) \equiv G_b(t)$$

where, as before, $\overline{S}_k$ denotes the inverse function of $\overline{B}_k$. Here, the bound is a little more complicated since the function $\overline{S}$ depends also on F_0. Inference about θ can be based on the following results.

Proposition 2 *Suppose F_0 belongs to a space of distribution functions as described in* (5.26)*. Moreover, suppose we have a random sample of size L of auctions each of which has $\mathcal{N}$ bidders and where we observe all bids. Take $k \in \mathbb{N}^+$, and let*

$$\begin{aligned} \Lambda(\theta|a,c;k) &= \int \Big(1 - \mathbb{1}\Big\{F_b(b) \ge F\big(\overline{S}_k(b;\mathcal{N}|\theta);\theta\big)\Big\}\Big)\mathbb{1}\Big\{a \le b \le c\Big\} dF_b(b), \\ \Gamma(\theta|k) &= \int\int \Lambda(\theta|a,c;k) dF_b(a) dF_b(c) \end{aligned} \tag{5.27}$$

Then, under the sole assumption that all bidders are level-k rational, the identified set is:

$$\Theta(k) = \Big\{\theta \in \Theta : \ \theta \in \underset{\theta \in \Theta}{argmin}\ \Gamma(\theta|k)^2\Big\}.$$

If the following condition holds for a known k_0, a stronger identification result can be obtained.

Assumption B1 Suppose there exists k_0 such that all bidders are level-k_0 rational and, with positive probability, bids are equal to the level-k_0 bounds. That is, suppose $\Pr\Big(b_i \leq \overline{B}_{k_0}\big(v_i;\mathcal{N}\big|\theta_0\big)\Big) = 1$, and $\Pr\Big(b_i = \overline{B}_{k_0}\big(v_i;\mathcal{N}\big|\theta_0\big)\Big) > 0$.

Proposition 3 *Suppose Assumption B1 holds and let $\Theta(k_0)$ be as defined in Proposition 2. For $\theta \in \Theta$ let*

$$\mathcal{F}^c(\theta) = \Big\{\theta' \in \Theta : \ \overline{B}_{k_0}(v_i;\mathcal{N}|\theta') < \overline{B}_{k_0}(v_i;\mathcal{N}|\theta) \ \ w.p.1.\Big\} \tag{5.28}$$

Then, the identified set is

$$\Theta_0^* = \Big\{\theta \in \Theta(k_0) : \ \nexists\ \theta' \in \Theta \quad such\ that \quad \theta' \in \mathcal{F}^c(\theta)\Big\}. \tag{5.29}$$

Consequently, if there exist $\overline{\theta} \in \Theta(k_0)$ such that

$$F(\cdot;\theta) < F(\cdot;\overline{\theta}) \quad for\ all \ \ \theta \in \Theta(k_0), \tag{5.30}$$

then $\Theta_0^* = \big\{\overline{\theta}\big\}$ and consequently, $\theta_0 = \overline{\theta}$.

Under Assumption B1, $\theta \notin \Theta(k_0)$ implies $\theta \neq \theta_0$ and $\theta \in \Theta(k_0)$ holds only if $\theta \notin \mathcal{F}^c(\theta_0)$. Suppose we have $\theta, \theta' \in \Theta(k_0)$ such that $\theta' \in \mathcal{F}^c(\theta)$. Then, it cannot be the case that $\theta = \theta_0$ because $\theta' \in \mathcal{F}^c(\theta_0)$ would imply $\theta' \notin \Theta(k_0)$. Thus, any such θ can be discarded as the true θ_0. To see the potential usefulness of this result, suppose $\mathcal{F}_v^{\Theta}$ is a space of exponentially distributed valuations, Assumption B1 holds and we find that the largest value of $\Theta(k_0)$ is $\overline{\theta} < \infty$. This would immediately imply $\theta_0 = \overline{\theta}$. Figure 10 illustrates this result for the exponential distribution. Inference about θ can be based on either of the previous results for a given k assumed to satisfy the corresponding assumptions by estimating $\Gamma(\theta|k)$ e.g, by

$$\widehat{\Gamma}(\theta|k) = \binom{L\mathcal{N}}{3}^{-1} \sum_{m<\ell<n} \Big(1 - \mathbb{1}\Big\{\widehat{F}_b(b_\ell) \geq F\big(\overline{S}_k(b_\ell;\mathcal{N}|\theta);\theta\big)\Big\}\Big)\mathbb{1}\Big\{b_m \leq b_\ell \leq b_n\Big\} \tag{5.31}$$

where $\widehat{F}_b(b) = \frac{1}{L}\sum_{\ell=1}^{L} \mathbb{1}\big\{b_\ell \leq b\big\}$ is the empirical cdf. For a given θ, the level-k rationalizable upper bound $\overline{S}_k(\cdot;\mathcal{N}|\theta)$ can be computed analytically for any $k \geq 1$. If Θ is a subset of a Euclidian space, then recent set inference methods can be used to obtain estimates of the identified set.

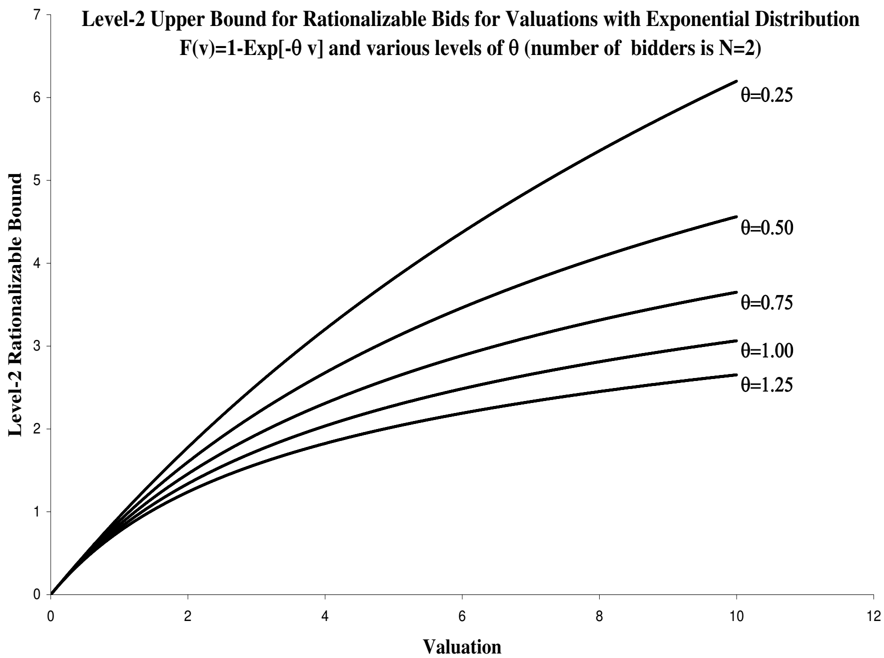


Figure 10: If Assumption B1 holds with $k_0 = 2$ and if we knew that $\{0.25, 0.50, 0.75, 1.00, 1.25\} \subset \Theta(k_0)$, it would follow immediately that $\theta_0 \geq 1.25$.

Remark 3 Let

$$\overline{B}_\infty(\cdot;\mathcal{N}|\theta) = \lim_{k\to\infty} \overline{B}_k(\cdot;\mathcal{N}|\theta). \tag{5.32}$$

Given our assumptions, the results in (BS) can be used to show that $\overline{B}_\infty(\cdot;\mathcal{N}|\theta)$ exists, and is a continuous, increasing, concave and invertible mapping that satisfies $\overline{B}_\infty(\cdot;\mathcal{N}|\theta) \geq b^{BNE}(\cdot;\mathcal{N}|\theta)$. If the rationality bound k_0 described in Assumption B1 does not exist, we must have $b_i \leq \overline{B}_\infty(v_i;\mathcal{N}|\theta)$ w.p.1. The results in Proposition 3 would follow if the conditions stated there hold for the mapping $\overline{B}_\infty(\cdot;\mathcal{N}|\theta)$.

5.2.1 Identification When Only Winning Bids are Observed

In this section we modify the results above for the case where only the wininng bid is observed in every auction. In particular, we observe

$$b^* = \max_{i=1,\ldots,\mathcal{N}} b_i. \tag{5.33}$$

Under these conditions, it follows from the monotonic nature of rationalizable upper bounds that if bidders are level-k rational, with probability one,

$$b^* \equiv \max_{i=1,\dots,\mathcal{N}} b_i \leq \max_{i=1,\dots,\mathcal{N}} \overline{B}_k(v_i;\mathcal{N}|\theta_0) = \overline{B}_k\big(\max_{i=1,\dots,\mathcal{N}} v_i;\mathcal{N}\big|\theta_0\big) \equiv \overline{B}_k(v^*;\mathcal{N}|\theta_0). \tag{5.34}$$

Then, we must have

$$\Pr(b^* \leq b) \geq \Pr\big(\overline{B}_k(v^*;\mathcal{N}|\theta_0) \leq b\big) \quad \forall\, b \in \mathbb{R}. \tag{5.35}$$

Since private values are iid, it follows that $v^* \sim F(\cdot;\theta_0)^{\mathcal{N}}$. Let $F_{b^*}(\cdot)$ denote the distribution function of b^*, the highest bid. Equation (5.35) becomes

$$F_{b^*}(b) \geq F\big(\overline{S}_k(b;\mathcal{N}|\theta_0);\theta_0\big)^{\mathcal{N}} \quad \forall\, b \in \mathbb{R}. \tag{5.36}$$

where, as before, $\overline{S}_k(\cdot;\mathcal{N}|\theta)$ denotes the inverse function of the upper bound $\overline{B}_k(\cdot;\mathcal{N}|\theta)$. Clearly, by the nondecreasing properties of distribution functions, (5.36) holds for all $b \in \mathbb{R}$ if and only if it holds for all $b \in \mathbb{S}(b^*)$ (the support of b^*). We conclude that this implies

$$F_{b^*}(b) \geq F\big(\overline{S}_k(b;\mathcal{N}|\theta_0);\theta_0\big)^{\mathcal{N}} \quad \forall\, b \in \mathbb{S}(b^*). \tag{5.37}$$

Equation (5.37) can be used, as in the previous to conduct inference on the set of consistent models. To do that, a similar objective function as the one in Proposition 2 above can be used. The results in Proposition 3 would also follow if Assumption B1 holds for b^*.[9]

5.3 Introducing a Binding Reserve Price

Suppose there is a nonzero reserve price p_0 set by the seller, and publicly observed by all potential buyers. We modify Assumption B0 accordingly as follows.

Assumption B0' Assume now that *all bidders expect any bid $b \geq p_0$ to win with strictly positive probability, and this is common knowledge.* The implication of this for submitted bids is that $b_i \geq p_0$ *if and only if* $v_i \geq p_0$. We restrict attention to beliefs that assign positive probability only to bidding functions that are increasing for all $v \geq p_0$ and are equal to p_0 for $v = p_0$. Formally, let $\mathcal{B}(p_0)$ denote the space of all Borel-measurable functions of the form

$$\Big\{b : [0,\omega) \to \mathbb{R}_+ :\ b(v) < p_0\ \forall\ v < p_0; \quad b(p_0) = p_0,\ \text{and for all } v > p_0\text{: } b(v) \leq v,\ \text{and}\ v > v' \ \Rightarrow\ b(v) > b(v')\Big\}.$$

[9] It appears that even if B1 is assumed to hold for all bids, we would have to explicitly assume that it holds for b^* because, with heterogeneous beliefs, it is no longer true that the highest bid corresponds to the highest valuation among potential bidders.

$$(5.38)$$

We will let $\mathcal{N}$ denote the number of potential bidders in the population and denote $\mathcal{B}_{-i}(p_0) = \mathcal{B}(p_0)^{\mathcal{N}-1}$. Beliefs for bidder i are probability distributions defined over a sigma-algebra $\Delta_{\mathcal{B}_{-i}}(p_0)$, where this sigma-algebra is such that singletons in $\mathcal{B}_{-i}$ are measurable (see footnote 6). As before, conjectures are defined as degenerate beliefs that assigns probability mass one to a singleton $\{b_j\}_{j\neq i} \in \mathcal{B}_{-i}$. We maintain the assumption that $F_0(\cdot)$ and $\mathcal{N}$ are common knowledge among potential bidders.

A consequence of a binding reserve price is that the number of potential bidders $\mathcal{N}$ may no longer be equal to the number of bidders who participate in the auction. Potential bidders with valuation $v_i < p_0$ will not submit a bid. Beliefs for valuations $v < p_0$ will be irrelevant for participating bidders, except for the fact that it is common knowledge that $v_j < p_0$ implies $b_j < p_0$ w.p.1 for all potential bidders. As in the case of zero reservation price, restricting attention to beliefs in $\mathcal{B}(p_0)$ will yield rationalizable upper bounds which also belong in $\mathcal{B}(p_0)$. It also rules out ties in the characterization of expected utility for bidders with valuation $v \geq p_0$ (the only ones who participate in the auction). As in the case of zero reservation price, restricting attention to beliefs in $\mathcal{B}(p_0)$ will imply that Bayesian-Nash equilibrium (BNE) optimal bids are always rationalizable.

Level-k Rationalizable Bids

The construction of rationalizable upper bounds will follow the same interim-rationalizability steps as in Subsection 5.1. Any bidder i with $v_i \geq p_0$ whose bids satisfy

$$b \leq v_i \quad \text{w.p.1.} \tag{5.39}$$

is called *level-1 rational*. Higher-rationality levels are characterized as before. The decision problem for any bidder i with $v_i \geq p_0$ can now be expressed as

$$\max_{b \geq p_0} \big(v_i - b\big)\widehat{\Pr}_i\Big[\max\Big\{p_0, \max_{j\neq i} b(v_j)\Big\} \leq b\Big], \tag{5.40}$$

where $\widehat{\Pr}_i(\cdot)$ denotes bidder i's subjective probability, derived from his beliefs and knowledge of $F_0(\cdot)$. The optimal bid for any assessment in $\mathcal{B}_{-i}(p_0)$ for any bidder with $v_i = p_0$ will always be $v_i = p_0$. Focusing on the case $v_i > p_0$, the most pessimistic assessment in $\mathcal{B}_{-i}(p_0)$ is given by the conjecture "$b(v_j) = v_j$ for all $j \neq i$ such that $v_j \geq p_0$". The optimal expected utility for this assessment is

$$\max_{b \geq p_0} \big(v_i - b\big) F_0(b)^{\mathcal{N}-1} \equiv \pi_2^*(v_i; \mathcal{N}, p_0), \tag{5.41}$$

which follows because $\widehat{\Pr}_i\big[\max\{p_0, \max_{j\neq i} v_j\} \leq b\big] = F_0(b)^{\mathcal{N}-1}\mathbb{1}\{b \geq p_0\}$ (recall that F_0, $\mathcal{N}$ and p_0 are common knowledge among bidders). Using the same arguments that followed Equation (5.21), level-2 rational bidders with $v_i \geq p_0$ must satisfy

$$p_0 \leq b \leq v_i - \pi_2^*(v_i; \mathcal{N}, p_0) \equiv \overline{B}_2(v_i; \mathcal{N}, p_0). \tag{5.42}$$

$\overline{B}_2(v_i; \mathcal{N}, p_0)$ is the level-1 rationalizable upper bound for all bidders with $v_i \geq p_0$. It is continuous, increasing and invertible for all $v_i \geq p_0$, with $\overline{B}_2(p_0; \mathcal{N}, p_0) = p_0$. In particular, the inverse function of $\overline{B}_2(\cdot; \mathcal{N}, p_0)$ is well-defined for all values and bids $\geq p_0$. As before, we will denote this inverse function by $\overline{S}_2(\cdot; \mathcal{N}, p_0)$. Note that, in general, (5.41) has corner solutions. That is, there exists a range of valuations $v_i > p_0$ such that $\pi_2^*(v_i; \mathcal{N}, p_0) = \big(v_i - p_0\big)F_0(p_0)^{\mathcal{N}-1}$. This, of course, will not impact the continuity, monotonicity and invertibility properties of the upper bound $\overline{B}_2(\cdot; \mathcal{N}, p_0)$ for values $v_i \geq p_0$. Nothing can be said about rationalizable upper bounds for $v_i < p_0$, except that they lie strictly beneath p_0. Bounds for such range of valuations are irrelevant for the optimal decision process of bidders. Proceeding iteratively, the *level-k bound for rationalizable bids* is given by

$$b_i \leq v_i - \underline{\pi}_k^*(v_i; \mathcal{N}, p_0) \equiv \overline{B}_k(v_i; \mathcal{N}, p_0), \quad \text{where} \quad \underline{\pi}_k^*(v_i; \mathcal{N}) = \max_{b \geq p_0} \big(v_i - b\big)F_0\big(\overline{S}_{k-1}(b; \mathcal{N}, p_0)\big)^{\mathcal{N}-1}, \tag{5.43}$$

and $\overline{S}_{k-1}(\cdot; \mathcal{N}, p_0)$ is the inverse function of $\overline{B}_{k-1}(\cdot; \mathcal{N}, p_0)$, well-defined for all values and bids $\geq p_0$.

5.3.1 Identification Observing only Winning Bids

If we replace Assumption B0 with B0', all the results in Subsection 5.2.1 hold with a binding reserve price for all $v_i \geq p_0$ and $b_i \geq p_0$. Consider a semiparametric setting as the one described in (5.26), where the distribution of valuations is allowed to depend on the publicly observed reserve price p_0

$$\mathcal{F}_v^{\Theta, p_0} = \Big\{F(\cdot; \theta, p_0) : \ \theta \in \Theta, \text{ and } F_0(\cdot; p_0) = F(\cdot; \theta_0, p_0) \text{ for some } \theta_0 \in \Theta\Big\} \tag{5.44}$$

Let $\overline{B}_k(\cdot; \mathcal{N}|\theta, p_0)$ denote the k-level upper bound for rationalizable bids that would be induced by a given distribution $F(\cdot; \theta, p_0) \in \mathcal{F}_v^{\Theta, p_0}$, let $\overline{B}_k(\cdot; \mathcal{N}|\theta, p_0)$ denote its inverse function. Let b^* denote the winning bid, and $F_{b^*}(\cdot; p_0)$ denote its distribution function (given p_0). Note that b^* is $\max_{i=1,\dots,\mathcal{N}} b_i$, truncated from below at p_0. *This automatic truncation ensures that the bounds in* (5.43) *are satisfied.* As we mentioned previously, bids below p_0 may not satisfy these bounds. If bidders are level-k rational, for any reserve price p_0 we must have

$$F_{b^*}(b;p_0) \geq F\big(\overline{S}_k(b;\mathcal{N}|\theta_0,p_0);\theta_0,p_0\big)^{\mathcal{N}} \quad \forall\, b \in \mathbb{S}(b^*|p_0), \tag{5.45}$$

where $\mathbb{S}(b^*|p_0)$ is the support of b^* given p_0. This result is the equivalent to Equation (5.37).

Proposition 4 *Suppose F_0 belongs to a space of distribution functions as described in* (5.44)*. Moreover, suppose we have a random sample of size L of auctions each of which has $\mathcal{N}$ bidders and where we only observe the winning bid in every auction. Let the reservation price p_0 be known. Define*

$$\Lambda(\theta|a,c;k,p_0) = \int \Big(1 - \mathbb{1}\Big\{F_{b^*}(b;p_0) \geq F\big(\overline{S}_k(b;\mathcal{N}|\theta,p_0);\theta,p_0\big)^{\mathcal{N}}\Big\}\Big)\mathbb{1}\big\{a \leq b \leq c\big\}dF_{b^*}(b;p_0),$$
$$\Gamma(\theta|k,p_0) = \int\int \Lambda(\theta|a,c;k,p_0)dF_{b^*}(a;p_0)dF_{b^*}(c;p_0). \tag{5.46}$$

Then, under the sole assumption that all bidders are level-k rational, the identified set is:

$$\Theta(k,p_0) = \Big\{\theta \in \Theta:\ \theta \in \underset{\theta\in\Theta}{\mathit{argmin}}\ \Gamma(\theta|k,p_0)^2\Big\}.$$

Now, suppose we assume that winning bids satisfy Assumption B1 for some k_0. For any $\theta \in \Theta$ let

$$\begin{aligned} \mathcal{F}^c(\theta,p_0) &= \Big\{\theta \in \Theta:\ \overline{B}_{k_0}(v_i;\mathcal{N}|\theta',p_0) < \overline{B}_{k_0}(v_i;\mathcal{N}|\theta,p_0)\ \textit{w.p.1}\Big\} \\ \Theta_0^*(p_0) &= \Big\{\theta \in \Theta(k_0,p_0):\ \nexists\ \theta' \in \Theta(k_0,p_0) \quad \textit{such that} \quad \theta' \in \mathcal{F}^c(\theta,p_0)\Big\} \end{aligned} \tag{5.47}$$

Then, the identified set is

$$\Theta_0^* = \Big\{\theta \in \Theta:\ \theta \in \Theta_0^*(p_0)\ \textit{w.p.1 (with respect to } p_0\textit{)}\Big\} \tag{5.48}$$

Note that the identification result in (5.48) requires that we explicitly assume that Assumption B1 holds for winning bids, see footnote 9. We now discuss briefly how the previous result could be constructive for estimation in the case of varying reserve price. Suppose we have a sample of $\ell = 1,\dots,L$ auctions with a population of potential bidders as described in the conditions leading to Proposition 4. Suppose the researcher observes the winning bids $\{b_\ell^*\}_{\ell=1}^L$, reservation prices $\{p_{0_\ell}\}_{\ell=1}^L$ and the number of *actual* entrants in each auction, $\{M_\ell\}_{\ell=1}^L$. We assume that the observed winning bids and reservation prices $\{b_\ell^*, p_{0_\ell}\}_{\ell=1}^L$ are iid draws from the same population. Once again, this implies that beliefs are random draws from the same distribution, and the data generating process for the reservation price is also the same across auctions. The latter would be trivially satisfied if the reservation price is constant across all auctions. We assume that the number of potential entrants $\mathcal{N}$ is the same for all auctions in the sample. Given our assumptions, this implies $I_\ell = \sum_{i=1}^{\mathcal{N}} \mathbb{1}\{v_i \geq p_{0_\ell}\}$. Therefore,

$$E[I_\ell] = \mathcal{N} E_{p_{0_\ell}}\big[F(p_{0_\ell};\theta_0, p_{0_\ell})\big] \;\Rightarrow\; \mathcal{N} = \frac{E[I_\ell]}{E_{p_{0_\ell}}\big[F(p_{0_\ell};\theta_0, p_{0_\ell})\big]}. \tag{5.49}$$

Let

$$\mathcal{N}(\theta) = \frac{E[I_\ell]}{E_{p_{0_\ell}}\big[F(p_{0_\ell};\theta, p_{0_\ell})\big]}; \quad \widehat{\mathcal{N}}(\theta) = \frac{\widehat{E}[I_\ell]}{\widehat{E}_{p_{0_\ell}}\big[F(p_{0_\ell};\theta, p_{0_\ell})\big]}. \tag{5.50}$$

The results in Proposition 4 can be constructive for estimation purposes if we use analog objects of the form

$$\begin{aligned} &\widehat{\Lambda}(\theta|a,c;k,p_0) = \\ &\qquad \frac{\sum_{\ell=1}^{L}\Big(1 - \mathbb{1}\Big\{\widehat{F}_{b^*}(b^*_\ell;p_0) \geq F\big(\overline{S}_k(b_\ell;\widehat{\mathcal{N}}(\theta)|\theta,p_0);\theta,p_0\big)^{\widehat{\mathcal{N}}(\theta)}\Big\}\Big)\mathbb{1}\{a \leq b^*_\ell \leq c\}K_h\big(p_{0_\ell} - p_0\big)}{\sum_{\ell=1}^{L} K_h\big(p_{0_\ell} - p_0\big)} \\ &\widehat{\Gamma}(\theta|k,p_0) = \frac{\sum_{m=1}^{L}\sum_{n=1}^{L}\widehat{\Lambda}(\theta|b^*_m, b^*_n;k,p_0)K_h\big(p_{0_m} - p_0\big)K_h\big(p_{0_n} - p_0\big)}{\sum_{m=1}^{L}\sum_{n=1}^{L} K_h\big(p_{0_m} - p_0\big)K_h\big(p_{0_n} - p_0\big)}, \end{aligned} \tag{5.51}$$

where $\widehat{F}_{b^*}(b;p_0) = \frac{\sum_{\ell=1}^{L}\mathbb{1}\{b^*_\ell \leq b\}K_h(p_{0_\ell}-p_0)}{\sum_{\ell=1}^{L}K_h(p_{0_\ell}-p_0)}$ and $K_h(u) = K(u/h)$ for some kernel $K(\cdot)$ and bandwidth sequence h. Recall that the relevant distribution is $F_{b^*}(\cdot;p_0)$, the distribution of $\max\limits_{i=1,\ldots,\mathcal{N}} b_i$, truncated at p_0, conditional on p_0.

5.3.2 Identification Results for the Rationality Level k_0 in Assumption B1.

Suppose we assume that there exists a finite $k_0 \geq 2$ that satisfies the conditions of Assumption B1 (otherwise, see Remark 3). The results in Proposition 3 are constructive when $k_0 \geq 2$ is *assumed* to be known. Naturally, one would be interested in having an identification result for both θ and k_0 simultaneously. For any k define

$$\widetilde{\Theta}(k) = \Big\{\theta \in \Theta : \; \Gamma(\theta|k)^2 = 0\Big\}, \tag{5.52}$$

where $\Gamma(\theta|k)$ is as defined in Proposition 2. Let $\Theta(k)$ be as defined there. Then, the following holds

(i) If $k \leq k_0$, then $\widetilde{\Theta}(k) = \Theta(k)$ and therefore $\theta_0 \in \widetilde{\Theta}(k)$.

(ii) If $k > k_0$, then $\theta_0 \notin \widetilde{\Theta}(k)$.

If $k_0 \geq 2$, it follows that $\theta_0 \in \Theta(2)$. Using this, along with (i)-(ii) we have the following result.

Proposition 5 *Let $\Theta(k)$ and $\Gamma(\theta|k)$ be as defined in Proposition 2. Define*

$$\underline{\Gamma}(k) = \min_{\theta \in \Theta(2)} \Gamma(\theta|k)^2. \tag{5.53}$$

Then, if Assumption B1 is satisfied with $k_0 \geq 2$, the following results hold

(i) $\underline{\Gamma}(k) = 0$ for all $k \leq k_0$. However, $\underline{\Gamma}(k) = 0$ does not imply $k \leq k_0$.

(ii) $\underline{\Gamma}(k) > 0$ implies $k > k_0$.

It follows from Proposition 5 that any k' such that $\underline{\Gamma}(k') > 0$ can be ruled out as the true k_0 described in Assumption B1, implying that *there is a subset of bidders who are strictly less than level-k' rational.* At the same time, the set $\{k \in \mathbb{N} : \ \underline{\Gamma}(k) = 0\}$ includes all $k \leq k_0$, but it also includes some values $k > k_0$.

6 Conclusion

This paper examined the identification power of equilibrium assumptions in three simple games. We replaced equilibrium with a form of rationality (interim rationalizability) which includes equilibrium as a special case, and compared the identified features of the game under rationality and under equilibrium. The games we studied are stylized versions of empirical models considered and applied in the literature and hence insights provided here can be carried over to those empirical frameworks. We do not advocate dropping the equilibrium assumptions from empirical work. But rather, the paper simply examines the question of what is the identifying power of equilibrium in these simple setups. For example, it is not clear that one would want to drop equilibrium in a first price auction since the underlying interim-rationalizability based model may not provide strong restrictions on the observed bids as they relate to the underlying valuations. Ultimately, the researcher faces the usual trade-off between robustness and predictive power and a balancing act guided by the economics of the particular application at hand needs to be done. In addition, we do not advocate either the use of rationalizability per se as the basis for strategic interaction. But, we do note that it has received extensive attention by game theorists (see for example Morris and Shin (2003), Dekel, Fudenberg, and Morris (2007) and the references cited therein). Moreover, rationalizable outcomes are ones that are closest to what a reasonable decision maker would use. Interim rationalizability allows us to incorporate the concept of higher order beliefs into the econometric analysis through what we defined here as rationality levels.

Some questions remain to be answered and we leave those for ongoing and future work. As far as the above results, the paper here is concerned with identification. A natural extension would be to study the statistical properties of estimators proposed above and apply those estimators in empirical examples. Another avenue of research is to extend some of the ideas above to dynamic setups. It is well known that inference in dynamic games is hard[10] The identification question is complicated due mostly to the complexity of the underlying economic model that is much richer. For example, the presence of multiple equilibria and beliefs off the equilibrium, where no data is available, are hard problems to deal with. In addition, the computational problem is large and involves solving complicated fixed point maps. It might be possible to examine the identification power of other strategic concepts that would be natural in dynamic settings such as the self-confirming equilibria of Fudenberg and Levine (1993). In addition to examining the robustness to equilibrium assumptions, these identification framework can be used to study whether inference under these different strategic frameworks, though less sharp, is more practically useful for applied researchers.

References

Aguiregabiria, V., and P. Mira (2004): "Sequential Estimation of Dynamic Games," Working Paper, forthcoming Econometrica.

Andrews, D., S. Berry, and P. Jia (2003): "On Placing Bounds on Parameters of Entry Games in the Presence of Multiple Equilibria," Working Paper.

Aradillas-Lopez, A. (2005): "Semiparametric Estimation of a Simultaneous Game with Incomplete Information," Working Paper, Princeton University.

[10] Recent important contributions to this field have recently been made by Aguiregabiria and Mira (2004) Bajari, Benkard, and Levin (2005) Pesendorfer and Schmidt-Dengler (2004) and Pakes, Ostrovsky, and Berry (2005). See also Berry and Tamer (2006).

BAJARI, P., L. BENKARD, AND J. LEVIN (2005): "Estimating Dynamic Models of Imperfect Competition," Working Paper, forthcoming Econometrica.

BAJARI, P., H. HONG, AND S. RYAN (2005): "Identification and Estimation of Discrete Games of Complete Information," Working Paper.

BATTIGALLI, P., AND M. SINISCHALCHI (2003): "Rationalizable Bidding in First-Price Auctions," *Games and Economic Behavior*, 45, 38–72.

BERNHEIM, D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52(4), 1007–1028.

BERRY, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

BERRY, S., AND E. TAMER (2006): "Identification in Models of Oligopoly Entry," Working Paper.

BJORN, P., AND Q. VUONG (1985): "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation," Caltech Working Paper 537.

BRESNAHAN, T., AND P. REISS (1991): "Entry and competition in concentrated markets," *Journal of Political Economy*, 99, 977–1009.

CILIBERTO, F., AND E. TAMER (2003): "Market Structure and Multiple Equilibria in Airline Markets," Working Paper.

COSTA-GOMES, M., AND V. CRAWFORD (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96(5), 1737–1745.

COSTA-GOMES, M., V. CRAWFORD, AND B. BROSETA (2001): "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69(5), 1193–1235.

CRAWFORD, V., AND N. IRIBERRI (2007a): "Level-k Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 75, 1721–1770.

——— (2007b): "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?," *Econometrica*, 75(6), 1721–1770.

Dekel, E., D. Fudenberg, and D. Levine (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282–303.

Dekel, E., D. Fudenberg, and S. Morris (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2.

Dekel, E., and A. Wolinsky (2003): "Rationalizable Outcomes of Large Private-Value First-Price Discrete Auctions," *Games and Economic Behavior*, 43(2), 175–188.

Fudenberg, D., and D. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61, 523–546.

Guerre, E., I. Perrigne, and Q. Vuong (1999): "Optimal Nonparametric Estimation of First-Price Auctions," Forthcoming, *Econometrica*.

Ho, T., C. Camerer, and K. Weigelt (1998): "Iterated Dominance and Iterated Best Response in Experimental 'p-Beauty Contests'," *American Economic Review*, 88(4), 947–969.

Honoré, B., and E. Tamer (2005): "Bounds on Parameters in Panel Dynamic Discrete Choice Models," Working Paper, forthcoming in Econometrica.

Morris, S., and H. Shin (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics. Proceedings of the Eighth World Congress of the Econometric Society*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky., pp. 56–114. Cambridge University Press.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85(5), 1313–1326.

Pakes, A., M. Ostrovsky, and S. Berry (2005): "Simple Estimators of the Parameters in Discrete Dynamic Games (with entry/exit)," Working Paper.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2005): "Moment Inequalities and Their Applications," Working Paper.

Pearce, D. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52(4), 1029–1050.

Pesendorfer, M., and P. Schmidt-Dengler (2004): "Identification and Estimation of Dynamic Games," Working Paper.

Seim, K. (2002): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," Working Paper, Stanford Business School.

Stahl, D., and P. Wilson (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10(1), 218–254.

Tamer, E. T. (2003): "Incomplete Bivariate Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70, 147–167.

Yildiz, M., and J. Weinstein (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75(2), 365–400.

7 Appendix

Proof of theorem 1: From our previous analysis, we know that both players are Level-1 rational if and only if with probability one in $\mathbb{S}(X)$,

$$
\begin{aligned}
\Pr\big(\varepsilon_1 > X_1'\beta_1, \varepsilon_2 > X_2'\beta_2\big|X\big) \le \Pr(Y_1=0,Y_2=0|X) \le \Pr\big(\varepsilon_1 > X_1'\beta_1+\alpha_1, \varepsilon_2 > X_2'\beta_2+\alpha_2\big|X\big)\\
\Pr\big(\varepsilon_1 \le X_1'\beta_1+\alpha_1, \varepsilon_2 > X_2'\beta_2\big|X\big) \le \Pr(Y_1=1,Y_2=0|X) \le \Pr\big(\varepsilon_1 \le X_1'\beta_1, \varepsilon_2 > X_2'\beta_2+\alpha_2\big|X\big)\\
\Pr\big(\varepsilon_1 > X_1'\beta_1, \varepsilon_2 \le X_2'\beta_2+\alpha_2\big|X\big) \le \Pr(Y_1=0,Y_2=1|X) \le \Pr\big(\varepsilon_1 > X_1'\beta_1+\alpha_1, \varepsilon_2 \le X_2'\beta_2\big|X\big)\\
\Pr\big(\varepsilon_1 \le X_1'\beta_1+\alpha_1, \varepsilon_2 \le X_2'\beta_2+\alpha_2\big|X\big) \le \Pr(Y_1=1,Y_2=1|X) \le \Pr\big(\varepsilon_1 \le X_1'\beta_1, \varepsilon_2 \le X_2'\beta_2\big|X\big)
\end{aligned}
\tag{7.54}
$$

We denote the true parameter value by θ_0. To prove part (a), take any $\widetilde{\beta}_1 \neq \beta_{1_0}$ such that $\widetilde{\beta}_{\ell,1} \neq \beta_{\ell,1_0}$. Given this and the support properties of $X_{\ell,1}$, for any scalar d we can observe either of the following two events with positive probability: (i) $X_1'\widetilde{\beta}_1 + d > X_1'\beta_{1_0}$ or (ii) $X_1'\widetilde{\beta}_1 < X_1'\beta_{1_0} + \alpha_{1_0}$. Take case (i) first, with $d = \alpha_1$ (arbitrary); if $\widetilde{\beta}_{2,\ell}\beta_{2,\ell_0} > 0$ we can make $\widetilde{\beta}_2'X_2 \to +\infty$ and $\beta_{2_0}'X_2 \to +\infty$. By Assumption (A1), this yields $\Pr(\varepsilon_1 \le X_1'\widetilde{\beta}_1 + \alpha_1, \varepsilon_2 \le X_2'\widetilde{\beta}_2 + \alpha_2|X) \to \Pr(\varepsilon_1 \le X_1'\widetilde{\beta}_1 + \alpha_1|X)$, and $\Pr(\varepsilon_1 \le X_1'\beta_{1_0}, \varepsilon_2 \le X_2'\beta_{2_0}|X) \to \Pr(\varepsilon_1 \le X_1'\beta_{1_0}|X) < \Pr(\varepsilon_1 \le X_1'\widetilde{\beta}_1 + \alpha_1|X)$. Therefore, with positive probability as X_2 explodes, $\Pr(\varepsilon_1 \le X_1'\widetilde{\beta}_1 + \alpha_1, \varepsilon_2 \le X_2'\widetilde{\beta}_2 + \alpha_2|X) > \Pr(\varepsilon_1 \le X_1'\beta_{1_0}, \varepsilon_2 \le X_2'\beta_{2_0}|X) > \Pr(Y_1 = 1, Y_2 = 1|X)$, which violates (7.54). If $\widetilde{\beta}_{2,\ell}\beta_{2,\ell_0} < 0$, the result is easier to obtain by making $\widetilde{\beta}_2'X_2 \to +\infty$ and $\beta_{2_0}'X_2 \to -\infty$. For case (ii), drive $\widetilde{\beta}_2'X_2 \to -\infty$ and $\beta_{2_0}'X_2 \to -\infty$ if $\widetilde{\beta}_2\beta_{2_0} > 0$, or $\widetilde{\beta}_2'X_2 \to -\infty$ and $\beta_{2_0}'X_2 \to +\infty$ if $\widetilde{\beta}_2\beta_{2_0} > 0$. In either case we eventually obtain $\Pr(\varepsilon_1 > X_1'\widetilde{\beta}_1, \varepsilon_2 > X_2'\widetilde{\beta}_2|X) > \Pr(\varepsilon_1 > X_1'\beta_{1_0} + \alpha_{1_0}, \varepsilon_2 > X_2'\beta_{2_0} + \alpha_{2_0}|X) > \Pr(Y_1 = 0, Y_2 = 0|X)$, which violates (7.54). This establishes identification of $\beta_{\ell,1}$, an analog proof shows that $\beta_{\ell,2}$ is identified, which proves part (a).

To establish part (b) focus on the worst-case scenario, and take $\widetilde{\theta} \neq \theta_0$ where $\widetilde{\beta}_{d,p} \neq \beta_{d,p_0}$, but $\widetilde{\beta}_{\ell,p} = \beta_{\ell,p_0}$ for $p = 1, 2$ —the parameters of the unbounded-support shifters are fixed at their true values–. Identification here must rely on the properties of $X_{d,p}$, the bounded-support shifters. The condition in the statement of the proposition ensures that (i) or (ii) (above) hold even if we fix $\widetilde{\beta}_{\ell,p} = \beta_{\ell,p_0}$. To complete the proof of (b) we proceed as in the previous paragraph (note that we now have $\widetilde{\beta}_{\ell,p}\beta_{\ell,p_0} > 0$). The case $\widetilde{\beta}_{d,p} \neq \beta_{d,p_0}$ and $\widetilde{\beta}_{\ell,p} \neq \beta_{\ell,p_0}$ is straightforward along the same lines. Now, onto part (c). Consider $\widetilde{\theta}$ that is equal to θ_0 element-by-element except for $\widetilde{\alpha}_1 \neq \alpha_{1_0}$ and recall that the parameter space of interest has $\alpha_p \le 0$. Clearly, none of the lower bounds in (7.54) evaluated at $\widetilde{\theta}$ will ever be larger than the corresponding upper bounds evaluated at θ_0, and none of the upper bounds evaluated at $\widetilde{\theta}$ will ever be smaller than the corresponding lower bounds evaluated at θ_0. Therefore, without further assumptions $\widetilde{\theta}$ and θ_0 are observationally equivalent and α_1 is not identified. The only way we can proceed is by adding more structure on $\Pr(Y_1, Y_2|X)$. We have $\Pr(\varepsilon_1 \le X_1'\beta_{1_0} + \alpha_{1_0}) \le \Pr(Y_1 = 1|X) \le \Pr(\varepsilon_1 \le X_1'\beta_{1_0})$, therefore $\Pr(\varepsilon_1 \le X_1'\beta_{1_0} + \widetilde{\alpha}_1) > \Pr(Y_1 = 1|X)$ only if $\widetilde{\alpha}_1 > \alpha_{1_0}$. Therefore, $\widetilde{\theta}$ can violate (7.54) only if $\widetilde{\alpha}_1 > \alpha_{1_0}$. For any such $\widetilde{\alpha}_1$, let $\Delta = \Pr(\varepsilon_1 \le X_1'\beta_{1_0} + \widetilde{\alpha}_1) - \Pr(\varepsilon_1 \le X_1'\beta_{1_0} + \alpha_{1_0}) > 0$.

By the assumption in part (c), there exists a subset $\mathcal{X}_1 \in \mathbb{S}(X_1)$ such that $\Pr(Y_1 = 1|X) < \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \alpha_{1_0}) + \Delta = \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \widetilde{\alpha}_1)$. Make $X_2'\beta_{2_0} \to +\infty$ and the lower bound on the fourth inequality in (7.54) will be violated. This establishes part (c). Any $\widetilde{\theta} \neq \theta_0$ where $\widetilde{\alpha}_p \neq \alpha_{p_0}$ and either $\widetilde{\beta}_{\ell,p} \neq \beta_{\ell,p_0}$ or $\widetilde{\beta}_{d,p} \neq \beta_{d,p_0}$ can be shown not to be observationally equivalent to θ_0 using the same arguments as in the previous paragraphs given the assumptions in parts (a) and (b). $\square$

Proof of Theorem 2

Suppose there exists a subset of realizations in $\overline{\mathcal{X}}_1^* \subset \mathcal{X}_1^*$ such that

$$X_1'\beta_1 + \Delta_1 + \alpha_1 > X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} \quad \forall\ X_1 \in \overline{\mathcal{X}}_1^*. \tag{7.55}$$

By continuity of the linear index, and of the distribution H_1, for any $X_1 \in \overline{\mathcal{X}}_1^*$ we can find a pair $0 \leq \overline{p}^L(X_1) < \overline{p}^U(X_1) \leq 1$ such that

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\overline{p}^L(X_1)\big) < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\overline{p}^U(X_1)\big). \tag{7.56}$$

To see why $\overline{p}^L(X_1)$ and $\overline{p}^U(X_1)$ exist, fix $\overline{p}^U(X_1) = 1$. By continuity, there exists a small enough $\delta > 0$ such that $\overline{p}^L(X_1) \geq 1 - \delta$ satisfies (7.56). If condition (4.15) in Theorem 2 holds, then there exists $\mathcal{W}_1^* \subset \mathbb{S}(W_1)$ such that[11]

$$\begin{aligned}
&\text{Min}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0})\big|\mathcal{I}_1\big]\Big\} \geq \overline{p}^L(X_1) \quad \forall\ W_1 \in \mathcal{W}_1^*\\
&\text{Max}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0})\big|\mathcal{I}_1\big]\Big\} \leq \overline{p}^U(X_1) \qquad \forall\ W_1 \in \mathcal{W}_1^*.
\end{aligned} \tag{7.57}$$

By definition, we have

$$E\big[H_2(X_2'\beta_2+\Delta_2+\alpha_2)\big|\mathcal{I}_1\big] = \pi_2^L(\theta|k=2;\mathcal{I}_1); \quad E\big[H_2(X_2'\beta_{2_0}+\Delta_{2_0})\big|\mathcal{I}_1\big] = \pi_2^U(\theta_0|k=2;\mathcal{I}_1). \tag{7.58}$$

Combining (7.57) and (7.58),

$$\pi_2^L(\theta|k=2;\mathcal{I}_1) \geq \overline{p}^L(X_1); \quad \pi_2^U(\theta_0|k=2;\mathcal{I}_1) \leq \overline{p}^U(X_1) \quad \forall\ W_1 \in \mathcal{W}_1^*. \tag{7.59}$$

Combining (7.56) and (7.59),

$$\begin{aligned}
H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k=2;\mathcal{I}_1)\big) &\leq H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\overline{p}^L(X_1)\big) < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\overline{p}^U(X_1)\big)\\
&\leq H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^U(\theta_0|k=2;\mathcal{I}_1)\big) \quad \forall\ W_1 \in \mathcal{W}_1^*.
\end{aligned}$$

[11] Note trivially that since $\alpha_p \leq 0$ everywhere in Θ, we have

$$\begin{aligned}
&\text{Min}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0})\big|\mathcal{I}_1\big]\Big\}\\
&\qquad \leq \text{Max}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0})\big|\mathcal{I}_1\big]\Big\} \quad \text{w.p.1.}
\end{aligned}$$

$$(7.60)$$

This would correspond to the case described in the first line of Equation (4.16). Next, suppose (7.55) does not hold but there exists a subset of realizations $\overline{X}_1^{**} \subset \mathcal{X}_1^*$ such that

$$X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} > X_1'\beta_1 + \Delta_1 + \alpha_1 \quad \forall\ X_1 \in \overline{\mathcal{X}}_1^{**}. \tag{7.61}$$

Repeating the same arguments as above exchanging θ and θ_0, we would arrive at the equivalent of (7.60), namely

$$H_1\big(X_1'\beta_{1_0}+\Delta_{1_0}+\alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) < H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) \quad \forall\ W_1 \in \mathcal{W}_1^{**}. \tag{7.62}$$

This would correspond to the case described in the second line of Equation (4.16). The last remaining possibility is that neither (4.16) nor (7.61) hold. In this case,

$$X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} = X_1'\beta_1 + \Delta_1 + \alpha_1 \quad \forall\ X_1 \in \mathcal{X}_1^*. \tag{7.63}$$

Since X_1 has full column rank in $\mathcal{X}_1^*$, (7.63) implies $\beta_{1_0} = \beta_1$ and $\Delta_{1_0} + \alpha_{1_0} = \Delta_1 + \alpha_1$. Since $\theta_1 \neq \theta_{1_0}$, we must have either

$$\Delta_1 > \Delta_{1_0}, \quad \text{or} \quad \Delta_1 < \Delta_{1_0}. \tag{7.64}$$

Suppose $\Delta_1 > \Delta_{1_0}$. This immediately yields $X_1'\beta_1 + \Delta_1 > X_1'\beta_{1_0} + \Delta_{1_0}$ for all $X_1 \in \mathcal{X}_1^*$. By continuity, we can find a pair $0 \leq \underline{p}^L(X_1) < \underline{p}^U(X_1) \leq 1$ such that

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\underline{p}^U(X_1)\big) > H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\underline{p}^L(X_1)\big). \tag{7.65}$$

To see why $\underline{p}^L(X_1)$ and $\underline{p}^U(X_1)$ exist, fix $\underline{p}^L(X_1) = 0$. By continuity, there exists a small enough $\delta > 0$ such that $\overline{p}^U(X_1) \leq \delta$ satisfies (7.65). If condition (4.15) in Theorem 2 holds, then there exists $\mathcal{W}_1^{***} \subset \mathbb{S}(W_1)$ such that

$$\begin{aligned}
&\text{Min}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0})\big|\mathcal{I}_1\big]\Big\} \geq \underline{p}^L(X_1) \quad \forall\ W_1 \in \mathcal{W}_1^{***}\\
&\text{Max}\Big\{E\big[H_2(X_2'\beta_2 + \Delta_2)\big|\mathcal{I}_1\big]\ ,\ E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0})\big|\mathcal{I}_1\big]\Big\} \leq \underline{p}^U(X_1) \qquad \forall\ W_1 \in \mathcal{W}_1^{***}.
\end{aligned} \tag{7.66}$$

Using the definitions of $\pi_2^L(\theta|k=2;\mathcal{I}_1)$ and $\pi_2^U(\theta_0|k=2;\mathcal{I}_1)$ (e.g, Equation 7.58), we obtain

$$\pi_2^U(\theta|k=2;\mathcal{I}_1) \leq \underline{p}^U(X_1); \quad \pi_2^L(\theta_0|k=2;\mathcal{I}_1) \geq \underline{p}^L(X_1) \quad \forall\ W_1 \in \mathcal{W}_1^{***}. \tag{7.67}$$

Using (7.65), this yields

$$H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) > H_1\big(X_1'\beta_{1_0}+\Delta_{1_0}+\alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) \quad \forall\ W_1 \in \mathcal{W}_1^{***}. \tag{7.68}$$

This corresponds to a case like the one described in the second line of Equation (4.16). If $\Delta_1 < \Delta_{1_0}$, the same arguments as above while exchanging θ with θ_0 would lead us to conclude that there exists a set $\mathcal{W}_1^{4*} \subset \mathbb{S}(W_1)$ such that

$$H_1\big(X_1'\beta_{1_0}+\Delta_{1_0}+\alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) > H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) \quad \forall\, W_1 \in \mathcal{W}_1^{4*}. \tag{7.69}$$

We have now established Equation (4.16) in Theorem 2 for the case $k = 2$. The cases $k > 2$ follow immediately from here by recalling the monotonic property of rationalizable bounds which says that, with probability one,

$$\begin{aligned} &H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^L(\theta|k+1;\mathcal{I}_1)\big) \leq H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^L(\theta|k;\mathcal{I}_1)\big) \quad \forall\, k \geq 1 \\ &H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k+1;\mathcal{I}_1)\big) \geq H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k;\mathcal{I}_1)\big) \quad \forall\, k \geq 1. \end{aligned}$$

To see why this implies that the rationalizable bounds for Player 1's conditional choice probabilities are disjoint with positive probability for all $k \geq 2$, recall that the Level-2 bounds are given by

$$\begin{aligned} &\Big[H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big)\ ,\ H_1\big(X_1'\beta_1+\Delta_1+\alpha_1\pi_2^L(\theta|k=2;\mathcal{I}_1)\big)\Big] \quad (\text{for } \theta) \\ &\Big[H_1\big(X_1'\beta_{1_0}+\Delta_{1_0}+\alpha_{1_0}\pi_2^U(\theta_0|k=2;\mathcal{I}_1)\big)\ ,\ H_1\big(X_1'\beta_{1_0}+\Delta_{1_0}+\alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big)\Big] \quad (\text{for } \theta_0). \end{aligned} \tag{7.70}$$

It follows from our results that the Level-2 rationalizable bounds for θ are disjoint from those of θ_0 with positive probability. Since the bounds for $k > 2$ are contained in those of $k = 2$ w.p.1, it follows immediately that these bounds are also disjoint for $k > 2$. It follows that if the population of Player 1 agents are at least Level-2 rational, any θ with $\theta_1 \neq \theta_{1_0}$ will produce Level-2 bounds that are violated with positive probability. Thus, no such θ can be observationally equivalent to one that has $\theta_1 = \theta_{1_0}$ and consequently, θ_{1_0} is identified. Naturally, if the same conditions of Theorem 2 hold when we exchange the subscripts "1" and "2", then θ_{2_0} will be identified. $\square$

Notes

[1]This is important since there is theoretical work that challenges the multiplicity issues that arise under rationalizability. For example, Yildiz and Weinstein (2007) show that for any rationalizable set of strategies in a given game, there is a local disturbance of that game where these are the unique rationalizable strategies. This ambiguity about what is the exact game that is being played is exactly the reason why it is important to study the identified features of a model in the presence of multiplicity.

[2]The model of demand and supply uses equilibrium to equate the quantity demanded with quantity supplied thus obtaining the classic simultaneous equation model. Other literatures in econometrics, like job search models and hedonic equilibrium models explicitly use equilibrium as a "moment condition."

[3]There might be ways to obtain sharper inference in these classes of games. This was pointed out to us by Francesca Molinari.